Forensic speaker comparison of Spanish twins and non-twin siblings:
A phonetic-acoustic analysis of formant trajectories in vocalic sequences, glottal source parameters and cepstral characteristics

Eugenia San Segundo Fernández

University of York, UK

Postdoctoral Research Assistant
Department of Language and Linguistic Science
University of York
UK

Research problem

From a forensic phonetic perspective, the voice characteristics of twin pairs and non-twin sibling pairs have frequently awoken special interest because these speakers represent extreme examples of similarity. Distinguishing their voices poses a well-recognized challenge in this discipline (e.g. Künzel, 2010), as can be concluded from the literature

review carried out for this thesis. Nevertheless, the *per se* research interest in these sibling pairs is deeply rooted in the nature-nurture dichotomy, which transcends any possible forensic application. Therefore, in this investigation a comparative study has been undertaken between genetically identical speakers (monozygotic twin pairs, *MZ*) and non-genetically identical speakers (dizygotic twins, *DZ*; non-twin siblings; and unrelated speakers) in an attempt to shed some light on a relevant but under-researched question: to which extend is our voice determined by our genes (*nature*) and to which extent is it due to behavioral, environmental or social aspects (*nurture*)?

Objective

The general objective of this thesis has been investigating the phonetic characteristics of three main speaker groups: MZ twins, DZ twins and non-twin siblings (24, 10 and 8 participants, respectively). Their phonetic-acoustic similarities have been studied by also taken into account a reference population of unrelated speakers (12 subjects). For the 54 male Spanish speakers (North-Central Peninsular variety) recorded *ad hoc* for this study, three different analyses were carried out. On the one hand, we labeled and analyzed the F1-F3 formant trajectories of 19 Spanish vocalic sequences. Secondly, several naturally sustained [e] tokens were extracted from the speakers' spontaneous vowel fillers and their glottal source characteristics were analyzed. These two approaches were complemented with an automatic speaker recognition analysis carried out with the software *Batvox*, based on cepstral parameters.

Hypothesis

The research hypotheses have been established according to what is known so far about twin and non-twin sibling pairs—their shared genetic endowment and the environmental influences possibly affecting their voice. Accepting the *equal environment assumption* traditionally associated to the classic twin method (Segal, 1990), MZ cotwins and DZ cotwins were expected to share the same environmental influences, while DZ twin pairs would share only half the genetic information than MZ twin pairs. Siblings would share the same genetic endowment as DZ twins but, on average, less environmental factors, mainly because of the age gap between them. Finally, unrelated speakers would share neither nature nor nurture.

Accordingly, five working hypotheses have been established for this thesis. Firstly, it was considered that a speaker's voice would be similar to itself, i.e. from one recording session to another. This assumption is made for all speaker types (H1). Secondly, accepting that MZ twin pairs are the most similar speakers that can exist (because of their shared genes and shared environmental influences), we hypothesized (H2) that MZ intra-pair comparisons would yield matching scores similar to those obtained in intra-speaker comparisons. The third hypothesis (H3) suggested that DZ intra-pair comparisons would yield relatively large matching scores but not as large as in the case of MZ twins. In the fourth hypothesis (H4), we stated that the intra-pair comparisons in the case of brothers would yield matching scores over the background baseline. That means that brothers should be more similar than unrelated speakers because they share 50% of their genes, exactly the same as DZ twins, and they usually have environmental influences in common, although to a lesser degree than DZ twins. Finally, we hypothesized (H5) that a background baseline should exist for the matching scores obtained by the unrelated speakers.

Methodology / Theoretical Framework

The first chapter describes the main current methodologies in Forensic Speaker Comparison (FSC). As a result of this review, it was concluded that adopting a hybrid perspective, which combines traditional and automatic analyses, is the most comprehensive approach to speaker comparison. For that reason, the three-folded approach of this thesis combines (a) traditional phonetic-acoustic parameters with (b) not only features but also techniques which are characteristic of automatic methods. In chapter three, the methodological details for carrying out this investigation are described. This includes a description of the main characteristics (age, dialect, etc.) of the recruited participants. An *ad hoc* corpus has been designed and collected for this thesis, including five speaking tasks and a vocal control technique. Some details about the recording procedure are also presented in the third chapter, such as the material and technical characteristics of the recordings, as well as the data collection set-up. Finally, a description follows of the likelihood-ratio approach within which the results of the different analyses are offered.

Results

All the parameters tested for this investigation have proved to be genetically conditioned —to a greater or lesser extent— since the hypothesized decreasing scale *MZ > DZ > non-twin siblings > unrelated speakers* was observed regardless of the analysis approach and for most speaker comparisons (the rare discordant results were thoroughly discussed in the corresponding analysis chapter). Therefore, the proposed parameters would be useful for comparing speech samples of known and unknown origin, as found in legal cases. Moreover, as different features were tested depending on the type of analysis conducted, we could indicate separately which parameters (or combination of parameters) were found more useful in the formant-trajectory analysis, on the one hand, and in the glottal-source study, on the other hand.

Future studies could explore the fusion possibilities of the three different systems tested for this investigation. The independence of glottal features from vocal-tract characteristics makes them specially promising for an improvement of an overall forensic system performance. Besides, future research focusing on twins' voices should pay more attention to the concept of *epigenetics*, which was briefly described in chapter two. We have continuously referred throughout this thesis to two basic forces which would intermingle to explain the (dis)similarities in twins and non-twins' voices, namely, genetic and environmental factors. The often-neglected third factor, i.e epigenetics —or the study of the alteration in the expression of specific genes caused by mechanisms other than changes in the underlying DNA sequence— could be behind the striking dissimilarities found for certain twin pairs.

References

Künzel, H. (2010). Automatic speaker recognition of identical twins. *The International Journal of Speech, Language and the Law*, 17(2), 251–277.

Segal, N. (1990). The importance of twin studies for individual differences research. *Journal of Counseling & Development*, 68(6), 612–622.