A forensic phonetic investigation into the speech patterns of identical and non-identical twins

Deborah Loakes

Research Assistant, Lecturer (The University of Melbourne)
Pronunciation Consultant (New Careers Australia)
School of Languages and Linguistics
Level 5 Arts Centre, The University of Melbourne
Parkville, Victoria, Australia 3010.

Awarding Institution: The University of Melbourne, Australia.
Date of Award: August 2006

keywords: speaker identification, australian English, twins' speech, similar-sounding speakers, likelihood ratios, phonetic variability

This investigation contributes to the field of forensic phonetics, especially forensic speaker identification (following approaches by Nolan 1983, Nolan and Oh 1996, Rose 2002). The concepts of speaker identity and phonetic variability are addressed by analysis of similar-sounding twin pairs. Most are identical twins who have had similar educational and environmental influences, and thus sources of between-speaker variation are reduced. The data are forensically realistic in that spontaneous conversational speech (both direct and telephone recordings), as well as non-contemporaneous data, are assessed. The investigation is concerned with Australian English, and also develops a greater understanding of Melbourne English as a unique variety.

The data comprise two corpora:
1) non-contemporaneous conversational speech samples produced by four pairs of twins from Melbourne aged between 18 and 20 (in conversation with the researcher, recorded at six-month intervals over a one year period); and

2) telephone recorded speech from five pairs of twins aged between 27 and 32 (in an interview with a medical researcher in an unrelated study).

These data were collected with assistance from The Australian Twin Registry and The Queensland Institute of Medical Research.

The thesis demonstrates that speaker variation is governed by a speaker's physical dimensions, and shows that the concept of learned variation, or ‘choice,’ also plays a major role in completing the picture. It shows that speakers do not have completely unique voices and can overlap considerably in terms of phonetic output. However, evidence of speaker-specific behaviour, even between identical twin pairs, was found.

In Chapter 1, the research environment in which this thesis is situated is described. Chapter 2 outlines reasons for investigating the speech patterns of twin pairs. Essentially, twins’ speech was analysed to explore the degree of difference between speakers who are as similar as possible, and because twins’ speech allows analysis of the limits of between-speaker variation. Data collection and method is described in Chapter 3. An important methodological approach used in this thesis is that same-segment vowel tokens were analysed, rather than strictly controlling phonetic context. That is, formants 1–4 of all lexically stressed Australian English vowels drawn from varying phonetic contexts (except preceding rhotic, lateral and nasal consonants) were measured. Results of the analysis are described in Chapters 4–10; in Chapters 4–9 the focus is the ‘Melbourne’ corpus, while in Chapter 10 the focus is the ‘telephone’ corpus.

An auditory analysis is presented in Chapter 4, and differences between the twins’ Melbourne accent and speech in other areas of Australia (primarily Sydney/New South Wales) are outlined. Auditorily evident differences between members of the twin pairs are also discussed in Chapter 4. For example differences in vowel realisation were found where some speakers had consistently more fronted vowels than their twins (also confirmed by acoustic analysis). Chapter 5 presents a comparison of the twins’ formant patterns (F1-F4). Some phonetic differences between members of twin pairs are highlighted, and the results also show that twins’ speech can be very similar, even in spontaneously produced conversational speech. In Chapter 6, the non-contemporaneous data is assessed. Results show that F3 is the most speaker-specific formant frequency range and also that lax vowels are more speaker-specific than tense vowels. In
Chapter 7, the data are re-analysed using a likelihood ratio (LR) approach. Results confirm that twins’ speech is much closer in F-pattern than pairs of speakers from the ‘general population’ (unrelated pairs in the corpus). The results in this chapter thus validate the usefulness of analysing twins’ speech in forensic speaker identification, as they demonstrate that twins’ speech does in fact facilitate analysis of the very lowest limits of between-speaker variation. In Chapter 8, differences in production of selected consonants between members of twin pairs as well as across the corpus, are assessed. Consistent frication of /k/ and /p/ is shown to be speaker-specific, and as such these are potentially useful variables for analysis in the forensic context. In particular 1) one is almost 7 times more likely to observe frication of /k/ (in any prosodic position) at a threshold of 30% or more in known and unknown samples if they come from the same speaker rather than different speakers, and 2) one is almost 7 times more likely to observe frication of /p/ (in any prosodic position) at a threshold of 20% or more in known and unknown samples if they come from the same speaker rather than different speakers.

In Chapter 9, combinations of vowel and consonant parameters are assessed (rather than measuring variables in isolation) and the best combinations for discriminating the twins’ speech are presented. Results in this chapter show that three different sets of parameters were equally useful for discriminating twin speakers. These sets of parameters involved different acoustic parameters (e.g. varying vowels and formants) combined with a full set of auditory parameters (/k/ and /p/ frication, and presence of nasal voice in vowels). This is useful forensically because all possible auditory and acoustic parameters might not necessarily be present in a forensic sample, yet analysis of differing combinations of parameters can be of equal value. While results in Chapter 9 show, again, that between-speaker variation is greater than within-speaker variation, they also show how much overlap can occur between speech samples produced by two different individuals.

In Chapter 10, the telephone data are analysed and the same type of vowel parameters are shown to be speaker-specific across this group as amongst the ‘Melbourne twins’. Results in Chapter 10 show that analysis of telephone-transmitted speech does not preclude differences being observed between members of twin pairs, and in fact a greater number of differences are observed between the pairs in this corpus. A possible reason for this is that the twins in the telephone corpus are older than the Melbourne twins, and as such have a greater number of non-shared experiences, leading them to accommodate more with speakers other than their twin. The thesis is tied together in Chapter 11, where results are summarised, forensic implications are discussed, and suggestions for future research are outlined.
One of the forensic implications of this research is that analysis of same-segment tokens drawn from differing phonetic contexts is a valid approach for forensic speaker identification. That is, even though vowels were measured from a multitude of lexical items, between-speaker variation was consistently greater than within-speaker variation. This was the case even when data were drawn from non-contemporaneous spontaneous conversational speech produced by similar-sounding twin pairs. This finding is important forensically, showing that loosely controlled parameters can be incorporated into a forensic speaker identification case when little data or common text is available for analysis. Predictably however, greater speaker-specificity is seen when vowels are drawn from controlled contexts.

Another important implication of the research is that Melbourne English was found to be a more distinct variety of Australian English than has been previously reported. Earlier studies on Melbourne English have used auditory analysis and reported differences from vowels produced by speakers from Sydney (for example Bradley and Bradley 1979, Bradley 1980). However, prior to this thesis, no acoustic analyses of vowels produced by Melbourne speakers existed to show the actual degree of difference between regions.

Overall, the thesis shows that forensically realistic phonetic differences are observable between speakers who are theoretically as similar as possible. Importantly, this thesis also shows that twins’ speech can be discriminated – however, it is much easier to discriminate unrelated pairs of speakers.

References


