A Forensic Phonetic Investigation into the Speech Patterns of Identical and Non-Identical Twins

Deborah Loakes
The University of Melbourne, Australia
dloakes@unimelb.edu.au

ABSTRACT

Three pairs of similar-sounding twin pairs are analysed from a forensic phonetic perspective. The questions 'does each person in the population have a measurably unique voice?' (re. Nolan & Oh 1996:39) and can perceptually similar voices be adequately distinguished (re. Rose e.g. 1996, 1999a, 1999b)? are answered positively. The results cannot disprove the assertion that each person in the population has a measurably unique voice, and show that these closely related similar-sounding speakers can be distinguished phonetically (by analysis of F4 in most vowels and particular aspects of consonant production).

1. INTRODUCTION

1.1 Background

This paper is a presentation of research in progress. The overall study is an investigation into the speech patterns of Australian English speaking identical (MZ) and non-identical (DZ) twin pairs. The research is situated in the field of forensic speaker identification and concerns specifically, aspects of individuality of the voice. This paper addresses degrees of within- and between-speaker variation as well as the notion of the idiolect.

For speaker identification purposes, it is important to study the nature of phonetic variation between similar-sounding speakers with the same accent (re. Rose e.g. 1999a,1999b). In broad terms, the theories are that speaker variation is: *organic*, meaning that people sound different because they have differently shaped vocal tracts; or that it is *learned* and that variation is caused by more direct means (Nolan and Oh 1996:39).

Analysing differences is imperative, because 'an essential and unavoidable element of the forensic phonetician's task involves … an appraisal of the significance of the differences found' which essentially means phrasing conclusions not in a manner that highlights similarities between two or more speech samples, but rather illustrating the degree of difference between them (French 1994:179). It needs to be determined whether phonetic differences between the samples are 'diagnostic of different speakers or simply the type of differences one would expect from the same person speaking on different occasions…' (French 1994:179).

The second issue that arises in this type of research is the notion of the idiolect. For this study, the idiolect is understood as 'variation within a language that is associated with individual speakers' (Burridge & Mulder 1998:302). Another useful and more explicit definition of the idiolect is:

> even within a narrowly defined dialect community, individuals will have their own preferred detailed pronunciations of particular words. The combination of number of such preferred alternative pronunciations yields an overall pronunciation which is idiosyncratic, [and] that is, an individual's idiolect (Nolan 1994:331).[1]

It can be said then, that particular features which cannot be predicted from the speaker's accent are idiosyncratic. Specifically, 'an idiosyncratic feature is one which cannot be correlated with group factors such as sex, age, regional origin, social status, health, etc.' (Brown 1982:16). Nolan points out that the idiolect is 'really only a hypothesis' – he states that it is not known whether every person really does have a unique voice, and also that it is probably unlikely to be the case (1994:332, also see Hollien 2002:7). Despite this, a fundamental assumption in forensic phonetics is that people can be identified by their voice (Nolan 1994:331).

1.2 Twins

MZ twins have identical vocal tract anatomy. They 'share the same genetic material as a result of the division of the maternal egg in early embryonic life' (Decoster, Van Guysel, Vercammen and Debruyne 2001:50). It can be said then, that MZ twins are identical physiologically as well as in appearance (Nolan and Oh 1996:40). The term used to refer to non-identical twins is DZ, which is an abbreviation of *dizygotic*, meaning two eggs. Unlike MZ twins, DZ twins have divergent vocal tract anatomy. They 'are genetically related to the same degree as are brothers and sisters' (Decoster *et. al.* 2001:50).

Investigation into the speech patterns of MZ twins is ideal for studying organic variation. For MZ twins who have

[1] Also see Loakes (2002:215).

ISBN 1-876346-48-5

grown up in the same environment, variation in their speech patterns can be 'fairly confidently' attributed to choice (Nolan and Oh 1996:40).

1.3 Aims of this paper

Following the issues outlined in 1.2, the questions being addressed in this paper are 'Does each person in the population have a measurably unique voice?' (re. Nolan & Oh 1996:39) and can perceptually similar voices be adequately distinguished (re. Rose e.g. 1996, 1999a, 1999b)?

2. METHODOLOGY

2.1 Corpus

Currently, the overall data-set comprises 9 pairs of MZ and DZ twin pairs recorded by the researcher, and 14 pairs of MZ and DZ twins recorded via telephone transmission for the purpose of an external, unrelated study.

C1 and C2 are 20 year old MZ twins who have a general to cultivated Australian English accent (see Harrington, Cox and Evans:1997). The speech data comprises conversational speech between an interviewer and the participants (who were recorded independently of each other).

A questionnaire was administered to the twin pairs in order to make comparisons between and across the pairs. As comparisons are only being made between one twin pair in this paper, only responses which are relevant will be discussed here.

In terms of auditory similarity, both twins reported 'often' being confused on the telephone which is a criteria for their being classed as having similar sounding voices (re. Rose & Duncan 1995, Rose 1999a). In addition, C1 reports in the interview that his parents often become confused in discriminating their voices.

Further information on the twins' background is that they shared the same education until 18 months previous to the recording session and live in the same house. They do not share the same friends, and never spend time together except when they are at home. Neither report having any major surgery on any part of the vocal tract, and neither suffer any disease or condition affecting speech.

2.2 Analysis

Analysis is both auditory and acoustic, focussing on differences between twin pairs and across the sample.

The speech of one pair of similar-sounding MZ twins (C1 & C2) has been analysed both acoustically and auditorily with the purpose of investigating phonetic differences in their speech patterns. The results of this analysis are then compared with previous results from a primarily auditory analysis of MZ and DZ twins (Pair A and Pair B respectively) (see Loakes, 2002).

In the auditory analysis, the focus is on complicated production of consonant sequences where variation between speakers can be expected, the consonant sequences /stɹ/, /tɹ/ and /tʃ/. In addition to this, any evident idiosyncratic patterns are addressed with respect to consonant production.

In the acoustic analysis, the focus is between-speaker differences in mean formant values for F1-F4. Particular attention is paid to vowels where little variation between speakers is expected, for example /ɜ/ and /ɔ/ (suggested by Nolan p.c. 2001).

The subjects were recorded in the phonetics laboratory at the University of Melbourne. The recordings were digitised at 22.05 kHz using ESPS/Waves+ on a Sun Ultra 1-140 workstation – in the case of C1 and C2 from digital audio tapes, and in the case of Pair A and Pair B from audio tapes which had been copied from the original recordings. Consonant and vowel segments were labelled using The EMU Speech Database System, version 1.2 (Cassidy and Harrington, 2001).

For the acoustic analysis, the focus was on vowels in strong syllables which do not precede a nasal consonant. Mean values for formants 1 and 2 were extracted from the mid-point of vowels in these environments in order to convert the data to a bark scale. This was performed using the statistical package R, version 1.4.1. One-way ANOVAs were also conducted using R, in order to assess degrees of difference between F1-F4 for all vowels between the speakers. Where significant differences were found between speakers, one-way ANOVAs were performed to assess within speaker variation.

3. RESULTS AND DISCUSSION

3.1 Auditory analysis

Turning first to the auditory analysis, no evident differences were found in the twins' articulation of the sequences /stɹ/, /tɹ/ and /tʃ/.

In terms of other parameters, there are a number of phonetic features which could be considered idiosyncratic occurring in both twins' speech. In these cases, within-speaker variation is higher than between speaker variation, which is problematic in terms of the notion of the idiolect discussed previously.

A labiodental realisation of /ɹ/ occurs commonly (but not categorically) after bilabial plosives. So for both C1 and C2 /ɹ/ → [ʋ] in voiced and voiceless environments.

Another idiosyncratic feature which commonly occurs in the speech samples of both twins is bilabial plosives being realised as bilabial fricatives in word initial and intervocalic environments. There are numerous examples of both /p/ → [ɸ] and /b/→ [β] in both twins' speech samples.

A distinguishing feature between the samples is in the realisation of voiceless velar plosives. Commonly (but not categorically) for C1 /k/ → [x] both intervocalically and word finally. Where /k/ is not realised as a fricative in these environments, it is realised with a strongly fricated offglide, as /k/ → [k^x]. This does not occur at all in C2's speech patterns. In terms of consonant production, this is the only distinguishing feature between the twins. Figure 1 shows examples of C1's varying realisations of /k/ mentioned above, /k/ → [x] and /k/ → [k^x].

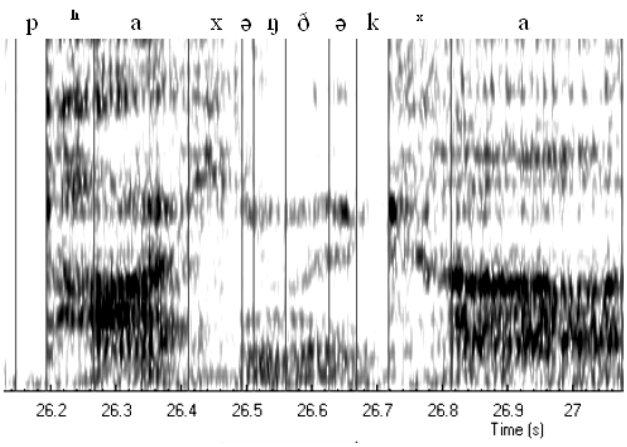

Figure 1: C1's realisation of the utterance *parking the car*.

3.2 Acoustic analysis

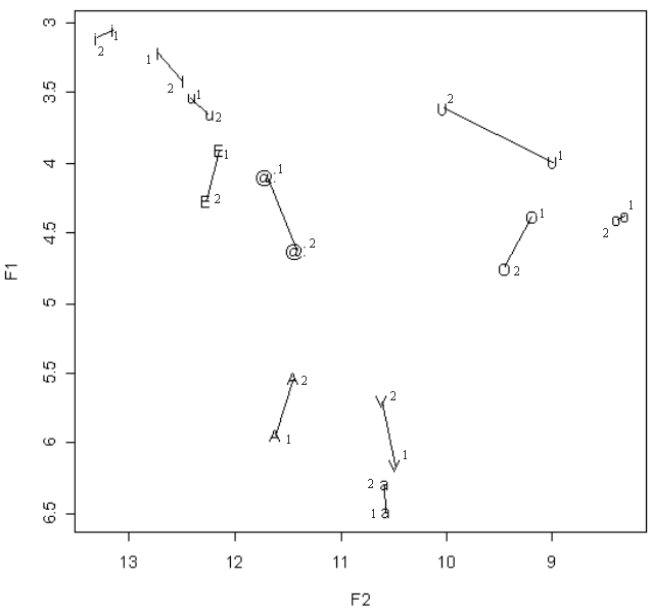

Figure 2: Bark scale comparison of C1 & C2's vowel space.

Figure 2 shows a bark scale comparison of formants 1 and 2 of C1 and C2's vowel space. The vowels represented in the figure are /i, ɪ, ɛ, æ, a, ɜ, ʌ, ɔ, o, ʊ, u/, which in machine readable phonetic alphabet correspond to /i, I, E, A, a, @:, V, o, O, U, u/ respectively. The distribution of vowels is typical for Australian English, including the close /u/, which is realised as [ʉ].

A visual comparison shows that as predicted, there is little variation between C1 and C2's realisation of /ɔ/. There is however, unexpected variation between their realisation of /ɜ/, with C1's realisation of the vowel being considerably closer than C2's.

From Figure 2, it can be said that C2's vowel space in terms of an F1 to F2 ratio is consistently more centralised than C1's. This however, is not statistically significant.

One-way ANOVAs (p=<0.05>) were performed to compare the formant values of F1-F4 for all the vowels under analysis.

In terms of F1, significant differences were between /æ/ (p=.01) and /ʌ/ (p=.02). No significant differences were found between F2 and F3 for any vowel.

For F4, significant differences were found between all the vowels except two. The degrees of difference between the speakers values for F4 are: /i/ p=0.004, /ɪ/ p= 0.008, /ɛ/ p=0.04, /æ/ p=0.01, /a/ p=0.005, /ɜ/ p=0.0001, /ʌ/ p=0.006, /ɒ/ p=0.05, /ɔ/ p=0.006. No significant differences were found between the twins' F4 values for /u/ and /ʊ/. In addition, no significant within-speaker differences were found between the F4 values within each speakers' data set.

3.3 Comparison with previous results

Auditory analysis of A1 & A2 (MZ twins) and B1 & B2 (DZ twins)[2] highlighted some phonetic differences between the pairs. The primary findings of this analysis were:

> For the identical twins, no systematic differences were found between the pair, but for the non-identical twins a [consistent] difference was found in the realisation of /u/. In addition … evidence was found for salient features which could not be predicted from the speakers' dialect actually occurring in the speech patterns of both twins in the pair.

Loakes (2002) also reports on a pilot study in which similar results were found in terms of idiosyncratic consonant production for similar sounding brothers, who

[2] See Loakes (2002).

ISBN 1-876346-48-5 © 2003 UAB

are less alike than twin pairs in terms of vocal tract anatomy. In this case though, more articulatory differences were found between the pair than in the cases involving twins.

In terms of auditory analysis the results of this paper confirm previous results. The acoustic analysis showed significant between-speaker differences for F4 on most vowels, and no significant within-speaker differences in the F4 of these vowels.

4. CONCLUSION

4.1 Concluding remarks

In this investigation, the auditory analysis served to uncover one consistent difference between the twin pairs in the realisation of /k/. The acoustic analysis showed that these twins can be distinguished in the value of F4 for stressed non-nasal vowels, except in the case of /u/ and /ʊ/.

Regarding the aims of the paper, both research questions have been answered positively. The results cannot disprove the assertion that each person in the population has a measurably unique voice. In addition, the results show that these closely related similar-sounding speakers can be distinguished phonetically. With respect to the findings on consonant production, it also appears that a narrower definition of the idiolect might be appropriate for undertaking speaker identification. The results suggest that while these speakers can be distinguished at the acoustic level particularly with regard to F4, the phonetic differences observed at the auditory level are minimal. Therefore the notion of the idiolect, while discernable at the acoustic level, might not always be upheld auditorily.

4.2 Future research

Further analysis will be undertaken to determine whether the results are similar between the other twin pairs in the data set, and to determine acoustically whether degrees of difference increase between similar sounding DZ twins and between similar sounding unrelated speakers. In addition a second recording session has been made of all the twin pairs in the data-set. This is in order to assess the validity of the results using non-contemporaneous speech samples (see Rose, e.g. 1999b, 2002). The Bayesian approach to forensic evidence will also be taken into account when analysing future results (re. Rose p.c., 2002).

ACKNOWLEDGEMENTS

The author would like to thank The Australian Twin Registry and The Queensland Institute of Medical Research.

REFERENCES

[1] R. Brown "What is speaker recognition" *Journal of the International Phonetics Association* **12 (1)**, pp. 13-24, 1982.

[2] K.Burridge and J. Mulder *English in Australia and New Zealand: An introduction to its structure, history and use* Melbourne: Oxford University Press, 1988.

[3] S. Cassidy and J. Harrington "Multilevel annotation in the EMU speech database management system" *Speech Communication* **37**, pp. 61-77, 2001.

[4] W. Decoster, A. Van Guysel, J. Vercammen, and F. Debruyne, "Voice Similarity in Identical Twins", *Acta oto-rhino-laryngologica Belgica* **55**, pp. 49-55, 2001.

[5] P. French "An overview of forensic phonetics with particular reference to speaker identification" *Forensic Linguistics: The International Journal of Speech Language and the Law* **1(2)**, pp.169-181, 1994.

[6] H. Hollien, H. *Forensic Voice Identification* San Diego: Academic Press, 2002.

[7] D. Loakes "Individuality of the voice: An analysis of telephone recorded twin pairs" *Proceedings of the 9th Australian International Conference on Speech Science & Technology* Canberra: Australian Speech Science and Technology Association, pp. 214-219, 2002.

[8] F. Nolan, "Auditory and acoustic analysis in speaker recognition" *Language and the Law* London: Longman, pp. 319-345, 1994.

[9] F. Nolan and T. Oh "Identical Twins, Different Voices", *Forensic Linguistics: The International Journal of Speech, Language and the Law* **3 (1)**, pp. 39-49, 1996.

[10] P. Rose and S. Duncan "Naïve auditory identification and discrimination of familiar listeners" *Forensic Linguistics: The International Journal of Speech Language and the Law* **2(1)**, pp.1-17, 1995.

[11] P. Rose "Differences and distinguishability in the acoustic characteristics of hello in voices of similar-sounding speakers: a forensic phonetic investigation" *Australian Review of Applied Linguistics* **22 (1)**, pp.1-42, 1999a.

[12] ___ "Long- and short-term within speaker differences in the formants of Australian hello" *Journal of the International Phonetic Association* **29 (1)**, pp. 1-31, 1999b.