A Forensic Phonetic Study of 'Dynamic' Sources of Variability in Speech: The DyViS Project

Francis Nolan, Kirsty McDougall, Gea de Jong and Toby Hudson
Department of Linguistics, University of Cambridge,
United Kingdom
fjn1@cam.ac.uk, kem37@cam.ac.uk, gd288@cam.ac.uk, toh22@cam.ac.uk

Abstract

DyViS is a forensic phonetics research project which investigates dynamic variability in speech from two perspectives: firstly, the use of dynamic (time-varying) properties of the speech signal for characterising speakers, especially formant dynamics; and secondly, the speaker-distinguishing potential of phonological variability due to sound change. In order to explore these two sources of dynamic variability, a large-scale database of speech material collected in simulated forensic conditions is being compiled. This paper outlines our motivation for the investigation, the methods and structure of the DyViS database, and some findings for the sound change study.

1. Introduction

The DyViS project ('Dynamic variability in speech: a forensic phonetic study of British English') at the University of Cambridge is a UK ESRC-funded research project [RES-000-23-1248] which aims to improve the forensic phonetic knowledge base. An increasing number of court cases involves the need to establish a speaker's identity from a recording of speech – a hoax emergency call, a fraudulent phone transaction, an obscene voicemail, etc. A phonetician is asked to carry out the task of 'forensic speaker identification', that is, to compare the speech on the incriminating recording with samples of speech from a suspect with a view to identifying the perpetrator or eliminating the suspect. Such analysis is constrained, however, by a conspicuous lack of knowledge about the distribution of speech features in the population as a whole. A phonetician is able to measure features of an individual's speech, but there is no known set of criteria that can determine speaker identity reliably. The task of characterising a speaker is complicated by the immense range of variation exhibited within the speech of an individual. In particular, speakers vary their voices depending on their familiarity with the interlocutor, their emotional state, the degree of formality of the situation, the level of background noise, and so on (Nolan, 1997: 748). A person's voice also changes with his or her state of health, and speakers can even disguise their voices. DyViS is developing a large-scale speech database designed for forensic phonetic research and the estimation of population statistics. The database is being used to investigate two kinds of dynamic variability in speech with respect to speaker characterisation. First, the reliability of a number of articulatory-acoustic dynamic features of speech as indices of speaker identity is being evaluated. Second, speaker identity is being explored in the context of the dynamics of diachronic sound change. This paper explains the rationale for investigating these two types of dynamic variability and outlines the structure of the DyViS database. It describes the techniques the DyViS project has developed for collecting partially phonetically controlled speech under simulated forensic conditions. Finally, some findings from the investigation of sound change are reported.

2. Quantifying dynamic features of speech encoding speaker identity: formant dynamics

The first type of 'dynamic variability' being investigated by DyViS concerns the rapidly changing parts of the speech signal known broadly as 'transitions'. Our assumption is that the speech signal contains linguistically determined targets (canonically thought of as the 'centres' of segments), linked by transitions. We hypothesise that the targets are highly constrained by the shared language system, and therefore there is greater scope for speaker-idiosyncrasy in the transitions, these being determined by the interaction of the specific organic endowment of the speaker, the adjacent linguistic targets, and the speaker's learned solution to moving between those targets.

DyViS builds on the recent doctoral research conducted by the second author (McDougall, 2005) which examined this idea for formant frequency dynamics (see also Greisbach, Esser and Weinstock, 1995; Ingram, Prandolini and Ong, 1996; McDougall,

Proceedings of the 11th Australian International Conference on Speech Science & Technology, ed. Paul Warren & Catherine I. Watson. ISBN 0 9581946 2 9
University of Auckland, New Zealand. December 6-8, 2006.

Accepted after full paper review

2004, 2006). Formant frequencies are highly relevant to speaker characterisation, since they are determined by both the dimensions of a speaker's vocal tract and the way the vocal organs are configured to produce each sound (Nolan, 2002: 78; Nolan and Grigoras, 2005). McDougall's experiments found that considerable speaker-specific information is present in the formant contours accompanying the transitions between the targets for a speech sound or sequence of sounds, in studies of F1-F3 of the sequence /aɪk/ in Australian English and of vowel-/r/-vowel sequences in British English. In these two studies, the degree of speaker discrimination achieved was a marked improvement on that attained by simply measuring formant frequencies at the temporal midpoint for each segment, which has been the typical approach in previous research.

Formant dynamics clearly provide a valuable source of speaker-specific information, however further research is needed to develop efficient ways of utilising this information. In the studies of /aɪk/ and /r/ the formant contours were essentially described with a series of instantaneous measurements. The next step is to develop a new technique to parameterise each formant curve such that the most defining aspects are captured for individual speakers with an economical descriptor (McDougall, 2005: ch. 6; see also McDougall, 2006).

DyViS aims to develop such a technique and to test it on a large population of speakers for a range of sound sequences. DyViS will also examine how well speaker-specific properties of formant dynamics are preserved across different recording sessions and under different conditions of linguistic context and speaking situation. This is very important for forensic speaker identification as the circumstances under which the incriminating and suspect samples are recorded are rarely the same (e.g. crime-related phone call versus police interview).

3. Testing diachronic change as a source of speaker idiosyncrasy

The second interpretation of dynamic variability in speech concerns linguistic change. Speech is not only dynamic in terms of transitions between sounds, but also in the sense that the language system is constantly in flux. Some linguistic variation leads to change as new realisations of existing contrasts become established, as old contrasts are subject to merger, and as new contrasts are formed. At any point in time, certain sounds are changing, while others appear more stable. This kind of dynamism is referred to in the DyViS project as 'diachronic dynamism'.

There are two ways in which this diachronic dynamism may be useful for identifying individual speakers. Firstly, particular speakers from the same social group may differ in terms of their realisations of variables which are undergoing change. One speaker may exhibit a more conservative or a more novel realisation than others. Although in the longer term a particular change would be expected to characterise all members of a speech community, in the shorter term this type of between-speaker variation might be valuable in distinguishing speakers (Moosmüller, 1997).

Furthermore, individual speakers might not use one particular realisation of a changing variable consistently, rather showing different rates of usage of novel versus conservative forms. For example, a speaker might use a particular realisation in a certain proportion of possible cases, or only in certain contexts, whereas others show different usage patterns (Loakes and McDougall, 2004). These usage rates could relate to the extent to which speakers have adopted variable patterns of phonological conditioning or lexical diffusion across the speech community as a whole. These patterns of usage might allow different individuals to be distinguished from one another (Butterfint, 2004).

DyViS will analyse the speaker-distinguishing potential of phonetic variables thought to be undergoing change in SSBE. This paper outlines the results of a comparison of the monophthongs /æ/, /ʊ/ and /uː/, which have been shown to be changing, with the relatively stable /iː/, /ɑː/ and /ɔː/ (Hawkins and Midgley, 2005).

4. The DyViS database

The DyViS project is developing a large-scale database of speech collected under simulated forensic conditions. The database will include recordings of 100 male speakers of Standard Southern British English (SSBE) aged 18-25 to exemplify a population of speakers of the same sex, age and accent group. Each speaker is recorded under both studio and telephone conditions, and in a number of speaking styles. The following tasks are undertaken by each subject:

1. simulated police interview (studio quality)
2. telephone conversation with 'accomplice' (studio and telephone quality)
3. reading passage (studio quality)
4. reading sentences (studio quality)

A subset of the speakers is participating in a second recording session to enable analysis of non-contemporaneous variation, since it has been suggested that non-contemporaneous samples will make speaker identification more difficult (Nolan, 1983: 12).

The speech files will be accompanied by orthographic labelling. Full orthographic transcripts of spontaneous data and lists of read sentences will also be provided. The database will be made publicly available at the conclusion of the project.

The simulated police interview is a new technique, devised specifically for DyViS, which it is hoped will be adopted and developed further by forensic phonetic

Proceedings of the 11th Australian International Conference on Speech Science & Technology, ed. Paul Warren & Catherine I. Watson. ISBN 0 9581946 2 9
University of Auckland, New Zealand. December 6-8, 2006.

Accepted after full paper review

researchers in the future. A pervading issue in forensic casework is the fact that the few databases that are available for determining population statistics are based on read rather than spontaneous speech, and in particular do not include spontaneous speech collected under 'stressful' circumstances as is frequently true of speech requiring analysis in forensic cases. DyViS aims to begin to remedy this situation by collecting spontaneous speech in a situation of 'cognitive conflict', where speakers are made to lie. The simulated police interview is an extension of the map task technique (Anderson *et al*., 1991) in which participants produce spontaneous conversation as prompted by the need to describe a route on a map. The names of features on the map are designed to include the target phonetic variables requiring elicitation.

In the police interview version of this task, the experimenter (here, the second or third author) assumes the role of police officer and the subject is the suspect being questioned. The experimenter and subject are seated facing each other at opposite ends of a table. On the table are two computer screens, back to back, linked to the same computer so that both participants are looking at the same display on each of the screens. The display is a PowerPoint presentation controlled by the experimenter with a mouse. Before the actual interview starts, the subject reads instructions on the screen explaining that he has taken part in the trafficking of heroin with another man, Robert Freeman, and that he (the subject) is being interviewed by the police. The subject is instructed that his memory and knowledge are represented in the maps and schemas he will see. His task is to be as co-operative as possible in answering the police officer's questions by using all information offered in black, but to avoid mentioning, or to deny, incriminating facts. All such facts are shown in red. For instance, he must deny knowledge of his accomplice, Robert Freeman. The subject is also told that it is okay to say 'I don't know' or 'I can't remember' when being interrogated about such information. The interview then commences and the experimenter/police officer guides the subject/suspect through the scenario represented on the slides, asking him to describe his whereabouts, actions and movements on the day in question, details of his friends and colleagues and so on. An example slide from the PowerPoint presentation is given in Figure 1.

The experimenter adapts the interview questions to ensure that all target items are elicited from the subject at least once. The experimenter also makes a point of pressing the subject about the information in red which he is not allowed to divulge. Although the set-up is not a completely realistic replication of a real police interview, the authors are confident that it does elicit spontaneous speech under relatively stressful circumstances; indeed a number of the subjects have commented afterwards that they found the task stressful.

Following the interview, the subject takes part in a telephone conversation with his 'accomplice'. The accomplice is played by the fourth author, who was a student at the same university as the subjects, and who happens to fall into the same age, sex and accent group as the subjects. It is thus hoped that the subjects will use a reasonably relaxed speaking style for this task, such as they might use when talking to a friend. The telephone call involves discussing what occurred in the police interview, so that the accomplice can 'get his story straight'. The subject is given cards showing the Power-Point slides from the interview for reference, and the accomplice talks the subject through the entire scenario, again taking care to ensure that each target item is elicited from the subject.

The third task is a reading passage in the form of a newspaper article about the same crime. The passage includes the same target words as the interview and telephone call to enable analysis of these items produced in reading style. The final task is reading a set of sentences which have been designed to elicit target variables in specific contexts in a reading/citation style.

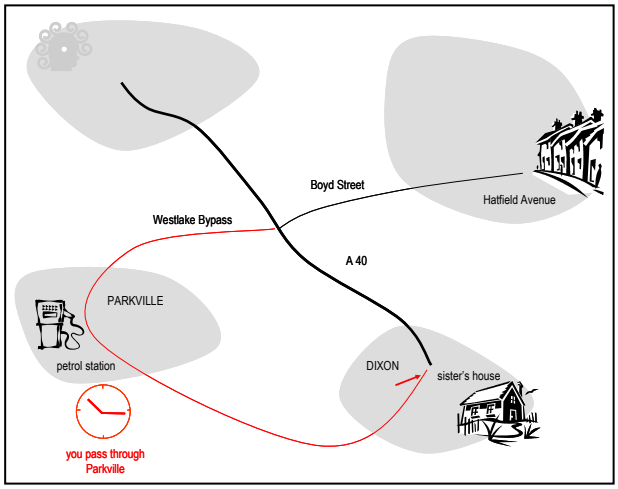

Figure 1. One of the slides from the police interview. The crime took place at the Parkville petrol station. The route taken by the suspect along Westlake Bypass and the time he passed through Parkville are shown in red (or grey for black and white print-outs) as he must not reveal taking this route at the time shown.

5. Diachronic sound change: SSBE monophthongs

The analysis reported here is of speech from the fourth task, the read sentences. Results for 20 subjects are presented, henceforth S1, S2, etc. The hypothesis under investigation is that vowels undergoing historical change (here /æ, ʊ, uː/ in SSBE) are likely to exhibit

Proceedings of the 11th Australian International Conference on Speech Science & Technology, ed. Paul Warren & Catherine I. Watson. ISBN 0 9581946 2 9
University of Auckland, New Zealand. December 6-8, 2006.

Accepted after full paper review

greater between-speaker variation than vowels which are relatively stable (SSBE /iː, ɑː, ɔː/).[1]

5.1. Materials and elicitation

The data analysed are six repetitions per speaker of the vowels /iː, æ, ɑː, ɔː, ʊ, uː/ in hVd contexts with nuclear stress. Each hVd word was included in capitals in a sentence, preceded by schwa and followed by *today*, as below:

> It's a warning we'd better HEED today.
> It's only one loaf, but it's all Peter HAD today.
> We worked rather HARD today.
> We built up quite a HOARD today.
> He insisted on wearing a HOOD today.
> He hates contracting words, but he said a WHO'D today.

Six instances of these sentences were arranged randomly among a number of other sentences. The sentences were presented to subjects for reading one at a time using PowerPoint. Subjects were asked to read aloud each sentence at a normal speed, in a normal, relaxed speaking style, emphasising the word in capitals. They practised reading the a few sentences at the start before the actual experimental items were recorded. Subjects were encouraged to take their time between sentences and asked to reread any sentences containing errors.

5.2. Recording

Subjects were recorded in the sound-treated booth in the Phonetics Laboratory in the Department of Linguistics, University of Cambridge. Each subject was seated with a Sennheiser ME64-K6 cardioid condenser microphone. The microphone was positioned about 20 cm from the subject's mouth. The recordings were made with a Marantz PMD670 portable solid state recorder using a sampling rate of 44.1 kHz.

5.3. Measurements

Analysis was carried out using *Praat* (www.praat.org). Wide-band spectrograms were produced for each utterance. Formant centre frequencies of F1 and F2 were generated by Praat's formant tracker to be written in a log file for each vowel at the time-slice judged by eye to be the centre of the vowel's steady state. In cases where no steady state for the vowel was apparent, the time-slice chosen was that considered to be the point at which the target for the vowel was achieved, according to movement of the F2 trajectory (i.e. a maximum or minimum or in the F2 frequency). All measurements were compared with visual estimates based on the spectrogram, values from adjacent time-slices, and the peak values of the spectral slice at that point. When values generated by *Praat* were found to be incorrect, they were replaced by correct values from a time-slice immediately preceding or following the slice being measured.

5.4. Results

The mean values of the frequencies of F1 and F2 of /iː, æ, ɑː, ɔː, ʊ, uː/ for each speaker are shown in the vowel quadrilateral plot in Figure 2. This figure shows that these vowels differ considerably from one another in the degree of between-speaker variation they exhibit. For example, /ɔː/ is tightly clustered in the vowel space, while /æ, ʊ, uː/ exhibit a wide range of realisations for different speakers. Consistent with the hypotheses based on patterns of sound change in SSBE, /ʊ/ and /uː/ demonstrate extensive variation in the F2 dimension and /æ/ varies widely in the F1 dimension. A result not predicted by sound change data for SSBE is that of considerable differences among speakers in their average F2 frequency of /iː/. /ɑː/ is also more variable in the F1 dimension than might be expected. However the formant frequencies are of course influenced by differences in vocal tract size as well as vowel quality; this issue is examined in more detail in the discussion section.

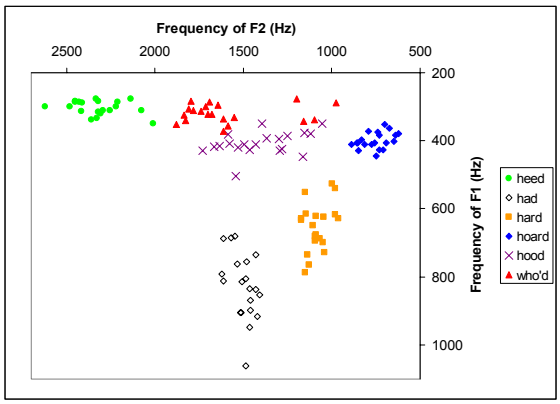

Figure 2. Mean F1 and F2 frequencies of /iː, æ, ɑː, ɔː, ʊ, uː/ for each speaker.

The degree of speaker-specificity exhibited by each vowel was tested using discriminant analysis. This analysis is a multivariate technique which can be used to determine whether a set of predictors can be combined to predict group membership (Tabachnick and Fidell, 1996: ch. 11). For the present study a 'group' is a speaker, or rather the set of utterances produced by a speaker. The discriminant analysis procedure constructs discriminant functions, each of which is a linear combination of the predictors that maximises differences between speakers relative to differences within speakers. These functions can be used to allocate each token in the data set to one of the speakers and determine a 'classification rate' according to the accuracy of the allocation. In the present study, this is done using the 'leave-one-out' method, where

[1] A more detailed report of the findings presented here will appear in de Jong *et al.* (forthcoming).

Proceedings of the 11th Australian International Conference on Speech Science & Technology, ed. Paul Warren & Catherine I. Watson. ISBN 0 9581946 2 9
University of Auckland, New Zealand. December 6-8, 2006.

Accepted after full paper review

each token is classified by discriminant functions derived from all tokens except for the token itself.

For each vowel quality, a direct discriminant function analysis was performed, using the F1 and F2 frequencies as predictors of 'membership' of the twenty different speakers, S1, S2, S3, … etc. ($k = 20$). The data set for each vowel contained the twenty speakers' six tokens, a total of 120 tokens. The resulting speaker classification rates are shown in Table 1.

Table 1: Speaker classification rates resulting from the discriminant analysis for each vowel. * indicates vowels changing in SSBE, according to previous research.

	Classification rate
HEED	35%
HAD*	35%
HARD	25%
HOARD	28%
HOOD*	41%
WHO'D*	27%

The discriminant analyses allocated the tokens to the correct speaker 25-41% of the time, rates much higher than chance (1/20 = 5%). However certain vowel qualities performed better than others. Reasons for the differing degrees of discrimination achieved are clearer when the data for individual speakers are examined, particularly with respect to within-speaker variation. For example, consider the different scenarios for /iː/, /ɔː/ and /uː/ represented by the six tokens of each vowel produced by five speakers shown in the F1-F2 plot in Figure 3. For /ɔː/ each speaker's tokens are clustered closely together in the vowel space. However for /uː/, some speakers produce very consistent realisations (S15 and S22) while others vary widely especially in the frequency of F2 (S2, S4 and S9). The situation for /iː/ is different again, with speakers exhibiting large between-speaker variation and small variation within-speaker.

Overall, for vowels where a speaker's average (F2, F1) realisation differs widely from one person to the next (Figure 2), but each individual is relatively consistent across his own productions, classification rates are higher. This is the situation for /iː/ (35%) and /ʊ/ (41%), especially due to the contribution of the F2 frequency. However vowel qualities which exhibit large within-speaker variation for certain speakers perform less well in the classification tests. This is the case for /uː/ (particularly due to F2) and /ɑː/ (particularly due to F1), with classification rates of 27% and 25% respectively. The vowel /æ/ with a rate of 35% also exhibits large within-speaker variation in the F1 direction, but this is compensated for by an extremely large between-speaker variation. /ɔː/, the most tightly clustered vowel in Figure 2, has low within-speaker variation, but its low between-speaker variation explains its lower classification rate of 28%.

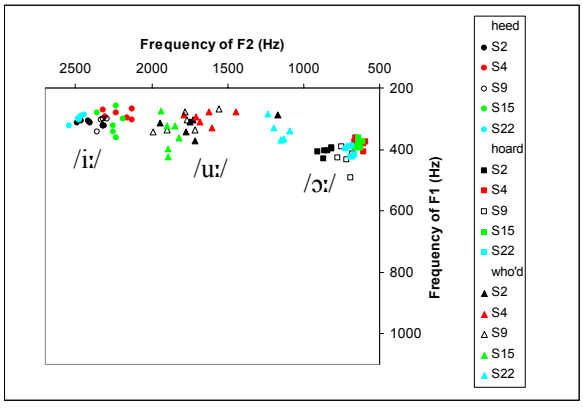

Figure 3. F1 and F2 frequencies of /iː/, /ɔː/ and /uː/ produced by S2, S4, S9, S15 and S22 (6 tokens each).

5.5. Discussion

The F1 and F2 frequencies of the vowels studied achieved differing levels of speaker discrimination. Patterns of sound change are relevant to the degree of speaker-specificity exhibited by a vowel; however vocal tract differences are also important, as explained below.

The only 'stable' vowel to yield a tight cluster of datapoints in F1-F2 space was /ɔː/; both between- and within-speaker variation were small for this vowel and the classification rate yielded was relatively low (28%). Large between-speaker variation was observed in the means for individual speakers for the F2 frequency of /iː/, /uː/ and /ʊ/ and the F1 frequency of /æ/ and to some extent /ɑː/ (see Figure 2).

The individual differences in the means for /uː/ and /ʊ/ are consistent with the predictions based on diachronic change, and this is confirmed by further auditory examination of tokens with F2 values at the extreme ends of the clusters for these vowels. However auditory examination of the data for /æ/ indicates that the differences in the frequency of F1 observed here do not correspond systematically to the degree of lowering of the vowel: some vowel tokens with similar coordinates do not necessarily sound the same in terms of vowel quality. While both vowel quality and vocal tract dimensions are conflated in the realisation of formant frequencies, it appears that for /æ/ vocal tract size plays a bigger role in explaining the differences among speakers for the F1 frequency than for other vowels. /æ/ and /ʊ/ yielded two of the highest classification rates on the discriminant analysis (35% and 41% respectively). However /uː/ performed less well (27%) due to the large within-speaker variation in this vowel for some speakers (see Figure 3). Although the data support the observation that /uː/ is changing in SSBE, this vowel did not perform as well in discriminant analysis, since for some individuals the diachronic change is reflected in considerable within-speaker instability.

The vowel /iː/ did not conform to the pattern of a tight cluster of data points expected for a 'stable' vowel,

Proceedings of the 11th Australian International Conference on Speech Science & Technology, ed. Paul Warren & Catherine I. Watson. ISBN 0 9581946 2 9
University of Auckland, New Zealand. December 6-8, 2006.

Accepted after full paper review

rather the vowel exhibited a wide spread in the F1 dimension. Nevertheless stability in /iː/ was evidenced by its low within-speaker variation, which, combined with the vowel's large between-speaker variation, led to a relatively high level of speaker discrimination (35% classification rate). Like /æ/, the individual differences in F2 for /iː/ are likely to be largely attributable to variation in vocal tract size rather than differences in auditory quality.

/ɑː/ was not as stable as /iː/ in terms of the range of F1 values exhibited by different speakers, although the cluster of datapoints for /ɑː/ was smaller than those of the three 'changing' vowels, and its classification rate was the lowest of the vowels tested (25%).

6. Conclusion

The DyViS project aims to improve the knowledge base in forensic phonetics by providing a large-scale database of speech collected under simulated forensic conditions to enable the estimation of population statistics for SSBE. In particular, a simulated police interview task has been developed, a novel technique for collecting phonetically-controlled speech in a situation of 'cognitive conflict'.

DyViS is investigating the relationship between speaker identity and dynamic variability in speech from two perspectives, that of acoustic-dynamic features of speech providing speaker-specific information, and that of 'diachronic dynamism' being a further source of differences among speakers.

This paper has provided an analysis of read data from the DyViS database relating to the diachronic aspect of the project, investigating whether sounds which are undergoing change are those most likely to differ among speakers. The results showed that this is true to an extent, but some qualifications are needed. The historically stable vowels of SSBE /ɔː/ and /ɑː/ offered the least reliable discrimination, and the rapidly fronting /ʊ/ provided the best discrimination. /uː/ is fronting, and was able to separate many speakers, but its within-speaker variation was large for some speakers, leading to a lower discrimination. /æ/ showed large between-speaker variation in F1, but this conflated slight audible phonetic quality variation with (inferred) vocal tract size differences. /iː/ was auditorily stable and provided good discrimination due to low within-speaker variation (presumably linked to both historical and proprioceptive stability) and large between-speaker variation (presumably related to vocal tract differences).

7. Acknowledgements

This research is supported by the UK Economic and Social Research Council [RES-000-23-1248]. We are also grateful to BT for sponsorship relating to the telephone transmission aspect of the project, to Mark J. Jones for his initial involvement in designing the project, and to Geoffrey Potter for technical assistance.

8. References

Anderson, A. H., Bader, M., Bard, E. G., Boyle, E., Doherty, G., Garrod, S., Isard, S., Kowtko, J., McAllister, J., Miller, J., Sotillo, C., Thompson, H. S., & Weinert, R. (1991). The HCRC Map Task Corpus. *Language and Speech*, 34 (4), 351-366.

Butterfint, Z. (2004). *Individuality in Phonetic Variation: An Investigation into the Role of Intra-Speaker Variation in Speaker Discrimination*. Ph.D. Dissertation, Manchester University.

de Jong, G., McDougall, K., & Nolan, F. (forthcoming). Sound change and speaker identity: an acoustic study. In C. Müller and S. Schötz (Eds.), *Speaker Classification*. Springer.

Greisbach, R., Esser, O., & Weinstock, C. (1995) Speaker identification by formant contours. In A. Braun and J.-P. Köster (Eds.), *Studies in Forensic Phonetics: Beiträge zur Phonetik und Linguistik*, *64* (pp. 49-55). Trier: Wissenschaftlicher Verlag Trier.

Hawkins, S., & Midgley, J. (2005). Formant frequencies of RP monophthongs in four age-groups of speakers. *Journal of the International Phonetic Association*, 35 (2), 183-199.

Ingram, J. C. L., Prandolini, R., & Ong, S. (1996). Formant trajectories as indices of phonetic variation for speaker identification. *Forensic Linguistics*, 3 (1), 129-145.

Loakes, D., & McDougall, K. (2004). Frication of /k/ and /p/ in Australian English: inter- and intra-speaker variation. In S. Cassidy, F. Cox, R. Mannell and S. Palethorpe (Eds.), *Proceedings of the 10th Australian International Conference on Speech Science and Technology* (pp. 171-176). 8-10 December 2004, Sydney: ASSTA.

McDougall, K. (2004). Speaker-specific formant dynamics: an experiment on Australian English /aɪ/. *International Journal of Speech, Language and the Law*, 11 (1), 103-130.

McDougall, K. (2005). *The Role of Formant Dynamics in Determining Speaker Identity*. PhD Dissertation, University of Cambridge.

McDougall, K. (2006). Dynamic features of speech and the characterisation of speakers: towards a new approach using formant frequencies. *International Journal of Speech, Language and the Law*, 13 (1), 89-126.

Moosmüller, S. (1997). Phonological variation in speaker identification. *Forensic Linguistics*, 4 (1), 29-47.

Nolan, F. (1983). *The Phonetic Bases of Speaker Recognition*. Cambridge: Cambridge University Press.

Nolan, F. (1997). Speaker recognition and forensic phonetics. In W. J. Hardcastle and J. Laver (Eds.), *The Handbook of Phonetic Sciences* (pp. 744-767). Cambridge: Cambridge University Press.

Nolan, F. (2002) The "telephone effect" on formants: a response. *Forensic Linguistics*, 9 (1), 74-82.

Nolan, F., & Grigoras, C. (2005). A case for formant analysis in forensic speaker identification. *International Journal of Speech, Language and the Law*, 2 (2), 143-173.

Tabachnick, B. G., & Fidell, L. S. (1996). *Using Multivariate Statistics*. New York: Harper Collins.

Proceedings of the 11th Australian International Conference on Speech Science & Technology, ed. Paul Warren & Catherine I. Watson. ISBN 0 9581946 2 9
University of Auckland, New Zealand. December 6-8, 2006.

Accepted after full paper review