ECONOMIC AND SOCIAL RESEARCH COUNCIL
END OF AWARD REPORT

For awards ending on or after 1 November 2009

This End of Award Report should be completed and submitted using the **grant reference** as the email subject, to **reportsofficer@esrc.ac.uk** on or before the due date.

The final instalment of the grant will not be paid until an End of Award Report is completed in full and accepted by ESRC.

Grant holders whose End of Award Report is overdue or incomplete will not be eligible for further ESRC funding until the Report is accepted. ESRC reserves the right to recover a sum of the expenditure incurred on the grant if the End of Award Report is overdue. (Please see Section 5 of the ESRC Research Funding Guide for details.)

Please refer to the Guidance notes when completing this End of Award Report.

Grant Reference	**RES-000-23-1248**		
Grant Title	**Dynamic Variability in Speech: A Forensic Phonetic Study of British English ('DyViS')**		
Grant Start Date	**01/10/2005**	**Total Amount Expended:**	**£ 402,942**
Grant End Date	**31/12/2009**		
Grant holding Institution	**University of Cambridge (Department of Linguistics)**		
Grant Holder	**Prof. Francis Nolan**		
Grant Holder's Contact Details	**Address**	**Email**	
	Dept. of Linguistics University of Cambridge Sidgwick Ave, CB3 9DA	**fjn1@cam.ac.uk**	
		Telephone	
		01223 335060	
Co-Investigators (as per project application):		**Institution**	
Dr Mark Jones		**University of Cambridge**	
Dr Kirsty McDougall		**University of Cambridge**	

1. NON-TECHNICAL SUMMARY

Please provide below a project summary written in non-technical language. The summary may be used by ESRC to publicise your work and should explain the aims and findings of the project. *[Max 250 words]*

Increasingly, legal cases involve the need to determine the speaker of some recorded speech. Voice samples, however, are not like fingerprints. A person's voice varies, depending for instance on speaking style, health, and many other factors. Despite this, there is a core of similiarities in an individual's speech which can serve to separate that person from others even with an identical accent; but researching how best to characterise a speaker has been hampered by the lack of data on voice variation in the population. This project has helped solve this problem by creating a carefully controlled database of 100 male speakers of the same accent speaking in different speaking styles, including over the telephone.

Within the project the database has already been used, for instance, to establish statistics on the pitch of voices within the population; to show how, even within the same accent, ongoing linguistic change is reflected in fine differences between individuals' vowel systems; to quantify between-speaker distinctions in the main resonances of vowels; and to reveal the way those resonances vary – reflecting the interaction of an individual's vocal tract, acquired habitual movements of the speech organs, and the language produced, in a way which serves to characterise a speaker. The database has also been used in a satellite project (ESRC RES-000-22-2582) exploring how listeners perceive voices, and how that perception is affected by the telephone. A pre-release version has also been used by the leading UK forensic phonetic consultancy to test an automatic speaker comparison system. [249/250]

2. PROJECT OVERVIEW

a) Objectives

Please state the aims and objectives of your project as outlined in your proposal to the ESRC. *[Max 200 words]*

The overall aim of the project was to provide a novel approach to the characterisation of speakers for forensic speaker identification. Specifically, its objectives were to:

#1 Test the practicality of a 'speaker-space' for distinguishing members of a large speaker population
#2 Quantify articulatory-acoustic dynamic features for individual speakers
#3 Test diachronic change as a source of speaker idiosyncrasy
#4 Make available a database for wider use by other researchers, forensic phonetic practitioners, and other interested persons.

(Objectives will be referred to as [#1] etc. in section (d))
[88/200]

b) Project Changes

Please describe any changes made to the original aims and objectives, and confirm that these were agreed with the ESRC. Please also detail any changes to the grant holder's institutional affiliation, project staffing or funding. *[Max 200 words]*

The overall aim and objectives of the 'DyViS' project (as it is known) remained unchanged.

Before the project started Dr Mark Jones, who was named as an RA, was awarded a British Academy Postdoctoral Research Fellowship. He therefore withdrew from active participation in the project (agreed by ESRC), but remained available for consultation as he stayed in Cambridge throughout. His RA post was awarded to Dr Gea de Jong, who had considerable experience in forensic phonetic casework, and who left a lectureship at City University to take up the post.

The other main deviation from plan was that Dr McDougall took maternity leave from 05/11/2007, returning to part-time working on 04/08/2008. As a consequence, the project was extended (with ESRC approval) to 4:3 years, ending in December 2009. A further complication was that Dr McDougall was a co-applicant on the satellite project RES-000-22-2582 'Voice similarity and the effect of the telephone', and 5% of her time and salary were allocated to that project in the calendar year 2008. This arrangement was mutually beneficial to the two projects, as the 'Voice similarity' project, rated 'Outstanding', was a casebook example of how the DyViS database can enable research in speaker identification. [199/200]

c) Methodology

Please describe the methodology that you employed in the project. Please also note any ethical issues that arose during the course of the work, the effects of this and any action taken. *[Max. 500 words]*

The general methodology was that of experimental phonetics, including the elicitation of controlled but naturalistic speech, and analysis using traditional (auditory) descriptive phonetics and acoustic analysis by computer.

A significant innovation was the method for eliciting speech. As well as two tasks which allowed for complete control over the words uttered, namely reading sentences and a continuous text, a new paradigm for interactional speech was employed. In Task 1, each subject played the role of a suspect being interviewed by the police [one of two RAs] about alleged drug dealing. During the interview, the currently relevant subset of facts was presented to the subject in one of a series of PowerPoint slides, for instance of a map showing a route taken. A few details on each slide would be in red; these represented incriminating facts which the subject must deny. The interaction was therefore similar to a police interview, with spontaneous speech and rapid thinking required to evade awkward questions, but with (by dint of placenames and other 'given' vocabulary) the experimenters retaining significant control over a subset of the words used, allowing for phonetic comparison. Task 2 was a 'friendly' debriefing telephone conversation with a third RA, playing an accomplice, during which further examples of these words were elicited. Task 2 provided for (a) comparison of telephone and interview styles, often relevant to forensic phonetic analysis, and (b) analysis and testing of the effects of telephone (landline) transmission, since the 'suspect's' side of the call was recorded simultaneously direct in the studio and over the public telephone network. The method yielded large amounts of natural spontaneous speech.

Auditory phonetic analysis was used to screen potential subjects, since only speakers with 'Standard Southern British' pronunciation were included in order to

achieve an accent-homogeneous database of 100 male speakers. It was also used as a preliminary stage in various analyses of the database, and the selection of speakers for the satellite 'Voice Similarity' project. Most acoustic analysis, and (orthographic) transcription, were carried out using Praat; scripts were employed for some tasks, such as the derivation of long-term fundamental frequency statistics. Because of the error-prone nature of automatic formant estimation, those studies (of the relation of the vowel system to historical change, and of formant dynamics as a cue to individual identity) which used formant frequencies measured them manually, with visual observation of formant patterns alongside LPC formant estimates. One study involving two other laboratories used the database to assess the replicability of such formant measurement, used in much forensic work (Duckworth et al., submitted to *IJSLL*).

Standard statistical tests were employed as appropriate. McDougall developed the use of polynomials to characterise formant dynamics, testing the speaker-discriminatory power of these by linear discriminant analysis. F0 and formants were compared to perceptual dimensions derived from Multidimensional Scaling of voice 'similarity' ratings.

No ethical issues arose. Tasks 1 and 2 were demanding, and potentially stressful, but no subjects experienced distress. Most seemed to have relished the challenge of spontaneous deceit! [496/500]

d) Project Findings

Please summarise the findings of the project, referring where appropriate to outputs recorded on *ESRC Society Today*. Any future research plans should also be identified. *[Max 500 words]*

* = details on ESRC SocietyToday

The successful collection of the database of recordings [#4], particularly the interactional tasks, demonstrated the effectiveness of the methodology outlined in section (c) above (see *Nolan et al. 2009, 'The DyViS database…', *IJSLL* 16(1)).

Work to quantify articulatory-acoustic dynamic features for individual speakers [#2] included analysis of formant dynamics in intervocalic /r/ sequences (*McDougall 2006 'The effects of stress…', BAAP) and an investigation of /uː / and /ai/ (e.g. *McDougall & Nolan 2007 'Discrimination of speakers…' *ICPhS2007*). Speaker-specific properties of formant dynamics in a casual speaking style as opposed to read speech have also been studied. A method capturing individuals' formant dynamics using polynomial equations was developed (cf. McDougall 2006 'Dynamic features of speech and the characterization of speakers…' *IJSLL* 13(1), 89–126). Papers reporting recent findings and further analyses are planned.

To test ongoing diachronic change [#3] as a source of speaker idiosyncrasy we investigated monophthongal vowels produced by the speakers in the database. Midpoint F1-F3 frequencies were measured for the vowels in (stable) HEED, HARD, HOARD, and (changing) HAD, HOOD, WHO'D. Results for subsets of up to 50 speakers are published e.g. in *de Jong et al. 2007 'Sound change and speaker identity…', *de Jong et al. 2007 'The speaker discriminating power…', *ICPhS2007*. Further analysis and articles for the complete set of measurements for 100 speakers are in preparation. Briefly, the historically stable vowel of HOARD has the lowest speaker-discrimination, but a simple picture for the others is complicated by instability *within* some speakers for the changing set, and physiologically-based differentiation (e.g. F1 of HARD, sensitive to pharynx length) in others of the stable set.

DyViS research aim [#1] was to investigate the notion of 'speaker space'. In addition to the vowel formant analysis above, we have analysed within-speaker variation in vowel formants (e.g. de Jong 2009 'Vowel space and within-speaker variability'

IAFPA2009). We have also presented f0 statistics for the full set of 100 speakers for the telephone call task (*Hudson et al. 2007 'F0 statistics for 100…' *ICPhS2007*) and for telephone speech in comparison with read speech (Hudson et al. 2009 'F0 trends… IAFPA2009). A journal article analysing f0 across a range of speaking styles is in progress, as is long term average formant analysis. Ultimately these will be combined in multidimensional speaker profiles.

Currently work is continuing on characterising the formant patterns in dynamic parts of the signal (McDougall & Wan, in prep, 'Individual variation in the formant dynamics of intervocalic /r/ in British English.') In the future a project is planned which would extend the database to other accents of English, in order to add that dimension of between-speaker differentiation to the personal characteristics explored in DyViS. This would be in the context of extending the work of the satellite project 'Voice similarity…', both to achieve a model of speaker similarity which would capture the relative weighting of personal and accent characteristics, and to provide the foundation for a more efficient and systematic strategy for the creation of voice parades. [495/500]

e) Contributions to wider ESRC initiatives (eg Research Programmes or Networks)

If your project was part of a wider ESRC initiative, please describe your contributions to the initiative's objectives and activities and note any effect on your project resulting from participation. *[Max. 200 words]*

The project was not part of a wider initiative, but it did enjoy a symbiotic relationship with our one-year ESRC funded project 'Voice similarity and the effect of the telephone' (RES-000-22-2582). The DyViS database provided the pool of accent-matched speakers from which samples could be selected for the voice similarity rating and voice parade experiments, and the rating judgments allowed correlations to be made between acoustic parameters measured in DyViS and dimensions emerging from Multidimensional Scaling of the ratings. [79/200]

3. EARLY AND ANTICIPATED IMPACTS

a) Summary of Impacts to date

Please summarise any impacts of the project to date, referring where appropriate to associated outputs recorded on *ESRC Society Today*. This should include both scientific impacts (relevant to the academic community) and economic and societal impacts (relevant to broader society). The impact can be relevant to any organisation, community or individual. *[Max. 400 words]*

Parts of the database have already been released to other researchers, notably J P French Associates in York, the leading UK firm of Forensic Phonetic analysts, for their research testing the BATVOX automatic speaker comparison system, Dr Angelika Braun (then) of the University of Marburg, for student projects on the phonetics of deception, and Martin Duckworth Consultancy for projects (involving Cambridge) (a) assessing replicability of formant measurements and (b) investigating speaker-specific properties of fluency disruptions. These applications all straddle fundamental research and practice (current or future) in forensic phonetic casework.

Our DyViS workshop 'Voices and Identity' (07/07/2008) drew professionals from fields such as law enforcement (e.g. detectives, investigators, scientific support officers), the law, and criminology. Its aim was to describe current practice and limitations in forensic phonetic casework, and to report relevant research developments from the DyViS project and its satellite VoiceSim [ESRC RES-000-22-2582]. The workshop was attended by 65 professionals from across the UK, and feedback was very positive. 'Knowledge Transfer' has also been achieved in individual events, e.g. Nolan addressed: the Criminal Bar Association (29/11/2008), a conference in Criminology (Cambridge) ('Evidence: possibilities and challenges, validity and value' – 24/09/2007), and a Forensic Human Identification course (Metropolitan Police Training Centre, Hendon 03/2007&2008); and spoke on Radio 4's 'Material World' (20/12/07). McDougall was interviewed on Channel 4 News about a prominent case involving speaker identity (02/10/2006), took part in Radio Cambridgeshire's 'Naked Scientists' popular science programme (12/10/2006), and talked to a joint meeting of Cambridge Medical Society and Cambridge District Law Society (21/01/2010). de Jong took part in an EU course for Turkish police officers (10/2007). All the above potentially have an impact on 'broader society', in that they facilitate progress to a greater understanding of phonetic evidence outside the domain of phonetics, and/or lead to enhancements in forensic phonetic methodology.

Academic impact has been achieved via the outputs listed on SocietyToday and reported synoptically above, specifically (to date) 18 conference papers, one book chapter, and 3 journal articles. The F0 statistics for 100 speakers are the first such large-scale data available for English, and constitutes an important benchmark. The detailing of the interaction between sound change and individual variation has implications for sociolinguistics as well as forensic phonetics, judging by the positive response to *McDougall & de Jong 2007 'Language change and the individual…' *UK-LVC*. Student research in York has taken up the 'formant dynamics' concept.
[396/400]

b) Anticipated/Potential Future Impacts

Please outline any anticipated or potential impacts (scientific or economic and societal) that you believe your project might have in future. *[Max. 200 words]*

We expect the DyViS database to enable research not only in forensic phonetics, but also language variation, conversation analysis, and the description of contemporary British pronunciation. It will, as mentioned above, continue to support research into a model of perceived voice similarity, and hence improved voice parades.

The critical mass of expertise and knowledge made possible by DyViS has already influenced forensic phonetic policy and practice, with the four DyViS researchers centrally involved in formulating the 'Position Statement concerning use of impressionistic likelihood terms in forensic speaker comparison cases' (*IJSLL* 14(1), 2007, 137–144) which has led to evidence being widely presented in a way which better reflects the limitations inherent in current speaker identification. Research carried out in DyViS, and the multiplicative effect of other researchers using the database, should mitigate those limitations and lead to methodological advances which inform a continuing review of practice.

Scientifically, the improved understanding of speaker characteristics emerging from DyViS may contribute to debates about how 'episodic' memory affects our perception of speech. Cues to the linguistic content of the speech signal are well known, but cues to the individual much less so; and yet we know that such information is attended to. [199/200]

You will be asked to complete an ESRC Impact Report 12 months after the end date of your award. The Impact Report will ask for details of any impacts that have arisen since the completion of the End of Award Report.

4. DECLARATIONS

Please ensure that sections A, B and C below are completed and signed by the appropriate individuals. The End of Award Report will not be accepted unless all sections are signed.

Please note hard copies are NOT required; electronic signatures are accepted and should be used.

A: To be completed by Grant Holder

Please read the following statements. Tick ONE statement under ii) and iii), then sign with an electronic signature at the end of the section.

i) The Project

This Report is an accurate overview of the project, its findings and impacts. All co-investigators named in the proposal to ESRC or appointed subsequently have seen and approved the Report.	☒

ii) Submissions to *ESRC Society Today*

Output and impact information has been submitted to *ESRC Society Today*. Details of any future outputs and impacts will be submitted as soon as they become available.	☒
OR	
This grant has not yet produced any outputs or impacts. Details of any future outputs and impacts will be submitted to *ESRC Society Today* as soon as they become available.	☐
OR	
This grant is not listed on *ESRC Society Today*.	☐

iii) Submission of Datasets

Datasets arising from this grant have been offered for deposit with the Economic and Social Data Service.	☒
OR	
Datasets that were anticipated in the grant proposal have not been produced and the Economic and Social Data Service has been notified.	☐
OR	
No datasets were proposed or produced from this grant.	☐