Front Vowels as Speaker-Specific: Some Evidence from Australian English

Deborah Loakes
Dept. of Linguistics and Applied Linguistics, The University of Melbourne, Australia
dloakes@unimelb.edu.au

Abstract

This investigation focuses on the degree of speaker-specificity in F2 and F3 of eleven Australian English monophthongs, determined using F-ratios from single-factor ANOVA. Non-contemporaneous samples of spontaneous speech produced by eight male speakers (four twin pairs) of Australian English were analysed, with the results indicating that the front vowels, and close-front vowels in particular, were most speaker-specific. The two most speaker-specific parameters, F2 and F3 of /ɪ/, were then used to compare same- and different-speaker pairs, with the results demonstrating between-speaker variation to be greater than within-speaker variation in the majority of cases.

1. Introduction

1.1. Speaker-specificity of Phonetic Parameters

This study investigates the degree of speaker-specificity of Australian English monophthongs in the speech patterns of a group of eight male speakers (four twin pairs) from Melbourne.

The term *speaker-specific* is used in this paper to describe a phonetic parameter that displays significantly greater between- than within-speaker variation. Analysis of such parameters is a requirement for forensic speaker identification, in which the aim is essentially to determine whether two or more speech samples were produced by the same speaker or by different speakers.

1.2. Previous Research

1.2.1. Speaker-Specificity of Vowels

Where vowels are concerned, previous research has revealed that parameters in the higher spectral region, such as F3 and F4, and sometimes F2 depending on the vowel, are more speaker-specific than those in the lower spectral region (Stevens, Carbonell and Woods 1968:1602, Sambur 1975; also see Lewis & Tuthill 1940, Ramishvili 1966, Dukiewicz 1970). More specifically, researchers have found that vowels which have a second formant in the higher spectral region, especially front vowels and close-front vowels in particular, are more useful than other parameters for highlighting speaker-specificity (Stevens *et. al*. 1968, Dukiewicz 1970, Pickett 2003).

Particular to Australian English spoken in Melbourne, Pickett (2003) analysed F2 and F3 of /ɛ æ ɔ ɜ/ produced by ten adult male speakers in 2 separate recording sessions of map-task elicited spontaneous speech. Pickett found that F3 of the front vowel /ɛ/ had the highest F-ratio in both the first and second recording sessions, though she did observe a degree of variability over time. She also found that both F2 and F3 of /ɜ/ fared well (Pickett 2003:47). Pickett's results for /ɛ/ are consistent with previous research indicating that front vowels are more speaker-specific than others, but the results for /ɜ/ appear less typical. However, assessing the overall mean formant frequencies of Pickett's data (2003: appendix 3.0), the realisation of /ɜ/ by the Melbourne English speakers is actually fronted. In fact for most speakers, the realisation of /ɜ/ is more front than /æ/ and almost as front as /ɛ/. Vowel spaces presented by Cox and Palethorpe for female adolescent speakers of Australian English from both regional Victoria and regional New South Wales show a similar distribution (2003:5).

1.2.2. Twins' Speech

The research presented in this paper is part of a wider investigation (Loakes, forthcoming) analysing the degree of between- and within-speaker differences in the speech of identical and non-identical twins; all who have similar-sounding voices and a shared environment, and, in the case of the identical twin pairs, who have the same shaped vocal tract anatomy (Decoster, Van Gysel, Vercammen & Debruyne 2001). This wider investigation is being undertaken to further knowledge regarding individuality of the voice for forensic speaker identification purposes, amongst pairs of speakers who are arguably as similar as possible.

Previous research investigating twins' speech has shown that while there are few differences between identical twins' acoustic output (Nolan & Oh 1996, Forrai & Gordos 1983) identical physical dimensions do not necessarily give rise to identical articulatory

Proceedings of the 10th Australian International Conference on Speech Science & Technology
Macquarie University, Sydney, December 8 to 10, 2004.

Accepted after full paper review

behaviour (Nolan & Oh 1996, Whiteside & Rixon 2003). That is, twins use differing articulatory strategies in approaching the same phonological targets (re. Nolan and Oh 1996).

1.3. The Current Investigation

In the current investigation, the focus is the metrically stressed monophthongal vowels of Australian English drawn from spontaneous speech. Because of the nature of the forensic context which typically involves telephone transmitted speech, in which the band-pass limits the frequencies available for analysis (Rose, Osanai & Kinoshita: 2003), only F2 and F3 are analysed in the current investigation for their relevance to the forensic context.

Rose (1999:38) called for forensic speaker identification research to analyse same-segment tokens such as stressed syllables, rather than one word, because in many forensic cases the same word is not available for comparison more than once. In addition, the number of tokens available for analysis in forensic speaker identification tasks are typically limited. As such, if speaker-specificity is demonstrated in vowel tokens which are not strictly controlled for phonetic context, the potential number of tokens available for analysis is improved.

In this investigation, speaker-specificity is measured by the use of F-ratios, which are a product of ANOVA and reflect both within- and between-sample variation (McCall 2001). From a forensic speaker identification perspective, a higher F-ratio reflects greater between- than within-speaker variation; so the higher F-ratio the more speaker-specific the parameter (Wolf 1972, Sambur 1975, Nolan 1983, Rose 2002).

The current investigation develops on previous research using F-ratios to determine the effectiveness of phonetic parameters for speaker identification by analysis of pairs of speakers who are as similar as possible, and by not strictly controlling for the phonetic environment of the vowel.[1]

1.3.1. Research Questions

The research questions for the current investigation are:

1. Which forensically realistic parameters, drawn from Australian English monophthongs, demonstrate the greatest speaker-specificity (as evidenced by high F-ratios)?
2. When comparing different speakers' speech samples, do the highest ranking parameters always result in high F-ratios?
3. When comparing one speaker's non-contemporaneous speech samples, do the highest ranking parameters always result in low F-ratios?

2. Method

2.1 Participants

The participants in this study are eight male speakers of Australian English from Melbourne (3 identical twin pairs and 1 non-identical twin pair), aged between 18 and 20. The speakers all have a shared environment; that is, each speaker was living with his twin at the time of the recording. In addition, the speakers are all university students, and while none are enrolled in the same degree, each speaker shared the same education as his twin until the end of high school (to 17 years of age).

The speakers in this investigation are referred to as TbY & TfY, PF & CF and LG & RG (identical twin pairs), and RH & ZH (non-identical twin pairs).

2.2 Recording, Labelling and Analysis

Each speaker took part in two Labovian-style interviews with the author, in which he discussed his interests and experiences. Approximately 8 minutes of spontaneous conversational speech from each interview (8 speakers x 2 sessions) is analysed in the present study. The participants were recorded in the phonetics laboratory at the University of Melbourne. The recordings were made on 120 minute Sony Digital Audio Tapes using a Sony ECM-999 PR electret condenser stereo microphone with a studio quality rack mount Tascam DA-30 DAT recorder.

Consonant and vowel segments were labelled using standard acoustic phonetic labelling procedures (c.f Croot & Taylor 1995) with The EMU Speech Database System, version 1.5.1 (cf. Cassidy and Harrington, 2001), using features from the waveform and wideband spectrogram. The vowels analysed in the current investigation are /i ɪ ɛ æ a ʌ ɒ ɔ ʊ u/. The analysis incorporated the mean centre formant frequencies of these vowels from F1 through F4, but only the results for the forensically realistic F2 and F3 are presented here. Only vowels which were in non-nasal metrically stressed environments were investigated, and vowels which preceded /l/ and /r/ were excluded from the analysis due to very strong coarticulatory effects. In diphthongised vowels, only the steady-state portion was analysed. Otherwise, the surrounding consonantal and vowel contexts of the target tokens were not strictly controlled.

To determine speaker-specificity, single factor ANOVA ($p = 0.05$) were undertaken with *R* version 1.4.1, to obtain F-ratios for each vowel and formant, incorporating both recording sessions for all eight

[1] However, strong coarticulatory effects have been controlled to a degree (see section 2.2).

Accepted after full paper review

speakers. The F-ratios for each parameter were ranked from most speaker-specific to least, and then single-factor ANOVA was used to compare each speaker with all others in the data-set for each of the two highest ranking parameters. This resulted in F-ratios for 28 different-speaker comparisons in both the F2 and F3 dimensions, allowing analysis of between-speaker variation (see section 3.3). Each speaker's non-contemporaneous speech samples were also compared using single-factor ANOVA, resulting in F-ratios for eight same-speaker comparisons in both the F2 and F3 dimensions, allowing analysis of non-contemporaneous within-speaker variation (see section 3.4).

3. Results and Discussion

3.1 Realisation of vowels

Before turning to the statistical results, this section discusses the realisation of some of the vowels in spontaneous speech by Melbourne speakers of Australian English, for whom acoustic descriptions of vowel spaces are rare (but see research by Bradley, e.g. Bradley & Bradley 1979). Particular vowels are discussed in cases where the realisation by the participants varies from standard descriptions of Australian English, or where the discussion will assist the understanding of the results.

Like Pickett's Melbourne English speakers, /ɜ/ is fronted by the speakers in this investigation and is more accurately [œ] in the majority of cases. F2 of /u/ varies substantially both within- and between-speakers, although for some speakers the realisation is relatively consistent. Amongst these participants, /u/ is often realised as a typical Australian English [ʉ], but most typically closer to [ʉ̟] and sometimes [y]. As an example of a typically fronted realisation, for one speaker (LG), the mean formant centre frequency observed for F2 of /u/ was 1975 Hz. A degree of within-speaker variation is also evident where F2 of /ʊ/ is concerned, so that for some speakers the vowel is realised variably, both as [ʊ̈] and [ʊ].

Other vowels are realised similarly to typical descriptions of Australian English. For example, /a/ and /ʌ/ are realised as [ɐ], with duration being the distinguishing feature between them (see Fletcher & McVeigh 1993), and /i/ is realised with variable degrees of schwa onglides. However, as mentioned above, only the steady-state of the phonological target has been measured in the present investigation.

3.2 Speaker-specificity of Aus. Eng. vowels

Turning now to the statistical results, Table 1 shows the ranking of parameters from most speaker-specific to least. Degrees of freedom are shown in the F-ratio column, and the actual F-ratio is presented in bold type.

The critical region in this analysis is 2.2. This is the level (expressed as the F-ratio) at which between-speaker variation is not significantly greater than within-speaker variation.

Table 1: F-ratios ranked from most speaker-specific to least

Rank	Parameter	F-ratio
1	F2 /ɪ/	F(15,695) = **22.4**
2	F3 /ɪ/	F(15,695) = **21.7**
3	F3 /æ/	F(15,650) = **20.1**
4	F3 /i/	F(15,669) = **17.4**
5	F2 /i/	F(15,669) = **16.04**
6	F3 /ɛ/	F(15,695) = **13.6**
7	F2 /ʌ/	F(15,632) = **13.4**
8	F2 /ɛ/	F(15, 695) = **12.4**
9	F3 /ɒ/	F(15,633) = **11.6**
10	F2 /æ/	F(15,650) = **11.4**
11	F2 /ɜ/	F(15,609) = **11.1**
12	F2 /u/	F(15,736) = **6.3**
13	F3 /u/	F(15,736) = **5.67**
14	F3 /ɜ/	F(15,609) = **5.5884**
15	F2 /ɒ/	F(15,633) = **5.5882**
16	F3 /a/	F(15,665) = **4.9**
17	F3 /ɔ/	F(15,639) = **4.4**
18	F3 /ʊ/	F(15,645) = **3.54**
19	F3 /ʌ/	F(15,632) = **3.1**
20	F2 /a/	F(15,665) = **2.21**
21	F2 /ɔ/	F(15,639) = **2.03**
22	F2 /ʊ/	F(15,645) = **0.66**

This table shows that of the 22 parameters, 20 display some degree of speaker-specificity. Overall, the highest ranking parameters are F2 and F3 of the close-front vowel /ɪ/, which have F-ratios of 22.4 and 21.7 respectively. This is followed by F3 of the front vowel /æ/ (20.1), F3 and F2 of the close front vowel /i/ (17.4 and 16.04) and F3 of the front vowel /ɛ/ (13.66). In other words, the six most speaker-specific parameters are from front vowels, and four of the five most speaker-specific parameters are from close-front vowels. This confirms the findings of previous research discussed in 1.2.1. In addition to these observations, it should also be noted that the three highest ranking parameters are lax vowels, which could be due to their short duration leaving less opportunity for within-speaker variability.

Proceedings of the 10th Australian International Conference on Speech Science & Technology
Macquarie University, Sydney, December 8 to 10, 2004. Copyright, Australian Speech Science & Technology Association Inc.

Accepted after full paper review

Parameters showing no speaker-specificity at all are F2 of the back vowel /ɔ/, and F2 of /ʊ/ which is realised between back and central. /ɔ/ is considered relatively invariable across speakers (Nolan p.c, 2001) and the low speaker-specificity of F2 of /ʊ/ is most likely due to the high degree of within-speaker variation for this parameter, discussed in section 3.1. While variation was also discussed for F2 of /ʉ/, this parameter ranks 12 of 22, indicating that while variation exists, between-speaker variation is still greater than within-speaker variation.

A possible reasons for greater speaker-specificity of front vowels is that they have both F2 and F3 values in the higher spectral region, which, as discussed in 1.2.1, is typically more speaker specific than the lower spectral region. In addition, this analysis involves three pairs of speakers who have vocal tracts with the same physical dimensions, and so between-speaker differences are not simply caused by differences in anatomy. Considering the way in which close-front segments are produced, with the tongue impeding the vocal tract more than for any other vowel (Peterson and Barney:1952), it is likely that individual differences in configuration of the tongue during articulation of close-front segments is responsible for their greater speaker-specificity observed in this analysis. In other words, speakers are using differing articulatory strategies in aiming to produce phonologically equivalent segments.

Differences between pairs of unrelated speakers and between twin pairs are investigated in the following section.

3.3 Comparing speakers

In this section, the focus is whether the highest ranking parameters are useful for discriminating different speakers. For the purposes of answering this research question, the focus is the two highest ranking parameters, F2 and F3 of /ɪ/. For this analysis, each speaker was compared with every other speaker in the data-set [2] using single-factor ANOVA, resulting in F-ratios for 28 different-speaker comparisons for each of F2 and F3 of /ɪ/. The results are presented in Table 2 below; asterisks indicate twin pairs and shading indicates instances in which the F-ratios were below the critical level, which in this analysis is 2.05.

From the table, it can be seen that in most cases that between-speaker variation is greater than within-speaker variation, though this is not categorical. In the F2 dimension, for six of the 28 cases, between-speaker variation was not shown to be greater than within-speaker variation. This is the case for only one of the twin pairs (the identical twins TbY & TfY for whom an F-ratio of 0.58 was observed); the comparison of all other twin pairs resulted in F-ratios higher than the critical level. In the F2 dimension, five of the six cases in which the parameters were not above the critical level involve the same combination of unrelated speakers; TbY, TfY, CF and ZH, and one involves the twins of two of these speakers (PF & RH).

Table 2: Between-speaker F-ratios: F2 & F3 /ɪ/

Pair	*df*	F-ratio F2	F-ratio F3
TbY_TfY*	(3, 176)	0.58	2.8
PF_CF*	(3, 170)	37.5	87.8
LG_RG*	(3, 173)	3.6	0.53
RH_ZH*	(3, 173)	7.9	6.6
TbY_PF	(3, 175)	19.7	28.4
TbY_CF	(3, 176)	0.1	0.3
TbY_LG	(3, 177)	25.6	13.2
TbY_RG	(3, 177)	7.2	5.1
TbY_RH	(3, 175)	7.7	10.6
TbY_ZH	(3, 179)	1.9	1.7
TfY_PF	(3, 175)	38.9	19.3
TfY_CF	(3, 171)	0.2	7.8
TfY_LG	(3, 173)	33.6	45.2
TfY_RG	(3, 173)	6.7	18.4
TfY_RH	(3, 170)	17.16	4.7
TfY_ZH	(3, 174)	2.1	0.1
PF_LG	(3, 172)	100.7	107.5
PF_RG	(3, 172)	73.8	73.7
PF_RH	(3, 170)	1.9	4.4
PF_ZH	(3, 174)	22.3	24.1
CF_LG	(3, 173)	30.1	16.1
CF_RG	(3, 173)	9.01	7.4
CF_RH	(3, 171)	14.9	36.1
CF_ZH	(3, 175)	0.8	6.2
LG_RH	(3, 172)	58.2	55.8
LG_ZH	(3, 176)	40.3	34.6
RG_RH	(3, 172)	35.5	36.9
RG_ZH	(3, 176)	12.9	15.9

In the F3 dimension, four of 28 F-ratios did not demonstrate between-speaker variation to be greater than within-speaker variation, and again this is the case for only one of the twin pairs (the non-identical twins LG & RG). In the other three cases, the same combination of speakers reported for F2 have F-ratios below the critical level, TbY, TfY, CF and ZH.

Overall, 22 of the 28 F-ratios in the F2 dimension, and 24 of the 28 F-ratios in the F3 dimension, are higher than the critical level of 2.05. Where the twin pairs are concerned, the F-ratios are mostly low; for example

[2] E.g. TbY&TfY, TbY&PF, TbY&CF etc.

Accepted after full paper review

where F2 of /ɪ/ is concerned, the comparison of LG & RH resulted in an F-ratio of 3.6 and the comparison of RH & ZH resulted in an F-ratio of 7.9. However PF and CF are an exception to this, with a high F-ratio in the F2 dimension (37.5) and an especially high F-ratio in the F3 dimension (87.8).

The highest F-ratios of all the comparisons were observed comparing PF and LG's speech samples, with F-ratios over 100 in both the F2 and F3 dimensions (100 and 107.5 respectively). The comparison of PF's speech samples with LG's twin also yielded relatively high F-ratios in both the F2 and F3 dimension (73.8 and 73.7 respectively).

These results show that the highest ranking parameters do not always result in high F-ratios when comparing different speakers. However, in most cases F-ratios are above the critical level, even when the analysis includes similar-sounding speakers (some with the same-shaped vocal tracts), spontaneous conversational speech, non-contemporaneous samples, and tokens which have not been controlled for consonantal context. Not surprisingly, smaller F-ratios have been observed when different-speaker analyses involve twin pairs, indicating greater acoustic similarity between the twin pairs than between unrelated speakers. However, in most cases involving twin pairs, some degree of speaker-specificity has demonstrated. That is, for all of the four twin pairs, at least one of F2 or F3 of /ɪ/ demonstrated a degree of speaker-specificity in having an F-ratio above the critical level.

3.4 Comparing different samples from one speaker

In this section, within-speaker variation is investigated by analysing F2 and F3 of /ɪ/ across each speaker's non-contemporaneous speech samples. For the parameters to be useful from a forensic speaker identification perspective results should be below the critical level, which in this analysis is 3.6.

The results are shown in Table 3 below; shading indicates that the F-ratio is above the critical level, and thus demonstrates significant within-speaker variation.

Taking into account both F2 and F3, in four of the 16 cases high F-ratios showed between-sample variation to be greater than within-sample variation. In these cases the F-ratios are *not* demonstrating speaker-specificity. The highest F-ratio was observed for TfY's F2 of /ɪ/ (13.1), demonstrating significant variation in this parameter across his speech samples. F-ratios above the critical level were also observed across TbY's, PF's and LG's speech samples for F3 of /ɪ/, with F-ratios of 9.8, 3.8 and 4.8 respectively.

Table 3: Within-Speaker F-ratios: F2 & F3 /ɪ/

Speaker	*df*	F-ratio F2	F-ratio F3
TbY	(1,90)	1.4	9.8
TfY	(1,85)	13.1	1.3
PF	(1,84)	5.12E-07	3.8
CF	(1,85)	3.2	0.05
LG	(1,86)	0.2	4.8
RG	(1,86)	0.5	0.4
RH	(1,84)	0.6	1.4
ZH	(1,88)	1.9	0.8

While this is the case, in most comparisons the F-ratios do not demonstrate significant within-speaker variation. That is, seven of the eight cases in the F2 dimension and five of the eight cases in the F3 dimension have F-ratios below the critical level. Most notably, a very small F-ratio has been observed for F2 of /ɪ/ across PF's speech samples (5.12E-07). It should also be noted that there are no speakers for whom within-speaker variation is significant across both the F2 and F3 dimensions. That is, in cases where an F-ratio is above the critical level for one formant, it is below the critical level for the other formant.

These results show that the highest ranking parameters do not always result in low F-ratios when comparing the same speaker's non-contemporaneous speech samples. However, F-ratios for F2 and F3 of /ɪ/ are below the critical level in the majority of cases, showing little within speaker-variation across a speaker's non-contemporaneous speech samples.

4. Conclusion

The results of this investigation have shown that front vowels in Australian English are more speaker-specific than other vowels. Likely reasons for this are that front vowels have both F2 and F3 in the higher spectral region, and also because the production of these vowels provides greater potential for individual behaviour in articulation strategies. In addition, greater speaker-specificity was observed in lax vowels as opposed to tense vowels, which is likely to be due to their shorter duration leaving less opportunity for within-speaker variability.

In further analysing the two most speaker-specific parameters, F2 and F3 of /ɪ/, between-speaker variation has been shown in most cases to be greater than within-speaker variation. The results have shown this to be the case even when dealing with similar-sounding twin speakers, spontaneous conversational speech, and without strictly controlling for consonantal context.

Proceedings of the 10th Australian International Conference on Speech Science & Technology
Macquarie University, Sydney, December 8 to 10, 2004.

Accepted after full paper review

The implication of these results is that F2 and F3 of /ɪ/ are also likely to be useful for demonstrating speaker-specificity in forensic speaker identification contexts involving Australian English. Individual differences in speech are likely to be even more evident when these parameters are analysed in conjunction with other front vowels, such as /i ɛ æ/, and with other segmental and suprasegmental parameters known to be useful for discriminating speakers.

5. Acknowledgements

Thanks to the Australian Twin Registry and the Queensland Institute of Medical Research for assistance in data collection.

6. References

Bradley, D. and M. Bradley (1979) 'Melbourne Vowels' *Working Papers in Linguistics* Vol.5, pp.64-84

Cassidy, S. and J. Harrington (2001) 'Multi-level annotation in the EMU speech database management system' *Speech Communication* (33), pp. 61-77

Cox, F. and Palethorpe, S. (2003) 'The Border Effect: Vowel Differences across the NSW-Victorian Border' *Proceedings of the 2003 Conference of the Australian Linguistic Society* <Cox & Palethorpe 2003> (2004, August 12)

Croot, K. and B. Taylor (1995) 'Criteria for Acoustic-Phonetic Segmentation and Word Labelling in the Australian National Database of Spoken Language' *Speech, Hearing and Language Research Centre Homepage*, Macquarie University <http://www.shlrc.mq.edu.au/andosl/> (2004, July 29)

Decoster, W., A. Van Gysel, J. Vercammen, and F. Debruyne (2001) 'Voice Similarity in Identical Twins' *Acta oto-rhino-laryngologica Belgica* Vol. 55., pp.49-55.

Dukiewicz L. (1970) 'Frequency Band Dependence of Speaker in Jassem, W. (ed.) Speech Analysis and Synthesis, Vol.II Polish Academy of Sciences, Warsaw, pp. 41-50.

Fletcher, J. and A. McVeigh (1993) 'Segment and Syllable Duration in Australian English' *Speech Communication* Vol.13, pp. 355-365

Forrai, G. and G. Gordos, (1983) 'A new acoustic method for the discrimination of monozygotic and dizygotic twins' *Acta Paediatrica Hungarica* Vol. 24 (4) pp. 315-321

Loakes, D. (forthcoming). *A Forensic Phonetic Investigation into the Speech Patterns of Identical and Non-Identical Twins* PhD Dissertation, University of Melbourne.

Lewis, D. and C. Tuthill (1940) 'Resonant Frequencies and Damping Constants of Resonators Involved in the Production of Sustained Vowels "O" and "Ah"' *Journal of the Acoustical Society of America* Vol.11, pp.451-456

McCall, R.B. (2001) *Fundamental Statistics for Behavioural Sciences (8th edition)* Wadsworth: California

Nolan, F. (1983) *The phonetic bases of speaker recognition* Cambridge: Cambridge University Press

Nolan, F. and T. Oh (1996) 'Identical twins, different voices' *Forensic Linguistics: The International Journal of Speech, Language and the Law* Vol. 3 (1) pp. 39-49

Peterson, G.E. and H.L. Barney 'Control Methods Used in the Study of Vowels' *Journal of the Acoustical Society of America* Vol.24(2), pp.175-184

Pickett, C. (2003) *Forensic Speaker Identification: An Acoustic Phonetic Study* Unpublished First Class Honours Thesis: Linguistics Program, Monash University

Ramishvili, G.S. (1966) 'Automatic Voice Recognition' *Engineering Cybernetics* Vol.5, pp.84-90

Rose, P. (2002) *Forensic Speaker Identification* Taylor & Francis: London

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

Rose, P., T. Osanai, and Y. Kinoshita (2003) 'Strength of Forensic Speaker Identification Evidence: Multispeaker Formant- and Cepstrum-Based Segmental Discrimination with a Bayesian Likelihood Ratio as Threshold' *The International Journal of Speech, Language and the Law: Forensic Linguistics* Vol. 10 (2) pp.179-202.

Sambur, M. R (1975) 'Selection of Acoustic Features for Speaker Identification' *IEEE Transactions on Acoustics, Speech, and Signal Processing* Vol.23 (2), pp.176-182

Stevens, K.N., C.E. Williams, J.R. Carbonell and B. Woods (1968) 'Speaker Authentication and Identification: A comparison of Spectrographic and Auditory Presentations of Speech Material' *Journal of the Acoustical Society of America* Vol.44 (No.6), pp.1596-1607

Stevens, K.N (2000) *Acoustic Phonetics* The MIT Press: Cambridge, Massachusetts

Whiteside, S.P. and E. Rixon (2003) 'Speech Characteristics of Monozygotic Twins and a Same-Sex Sibling: An Acoustic Case Study of Coarticulation Patterns in Read Speech' *Phonetica* Vol.60, pp.273-297

Wolf, J.J. (1972) 'Efficient Acoustic Parameters for Speaker Recognition' *Journal of the Acoustical Society of America* Vol. 51 (No.6, part 2), pp.2044-2056

Accepted after full paper review