

Contents

<i>List of tables</i>	viii
<i>List of figures</i>	ix
<i>Acknowledgments</i>	x
<i>Introduction</i>	1
1 Perspectives on speaker recognition	5
1.1 Speaker recognition	5
2 The bases of between-speaker differences	26
2.1 The received model	26
2.2 The inadequacies of the organic-learned dichotomy	27
2.3 Model of the sources of between-speaker differences	29
2.4 Summary and discussion	71
3 Short term parameters: segments and coarticulation	74
3.1 Introduction and survey	74
3.2 The experiments: aims and structure	77
3.3 Spectrograph analysis	85
3.4 Computer analysis	102
3.5 Conclusions	115
4 Long term quality	121
4.1 Introduction	121
4.2 The long term suprasegmental strand	122
4.3 The long term segmental strand	130
4.4 Acoustic correlates of supralaryngeal qualities	155
4.5 Long term properties: conclusions	192
5 Conclusions	197
5.1 Summary	197
5.2 Criticism of principle and practice in speaker recognition	198
5.3 Directions for speaker recognition	206
<i>References</i>	210
<i>Index</i>	219

Tables

3.1	Frequency of occurrence of English consonant phonemes	81
3.2	Word inventory	83
3.3	Overall means (all speakers) of formants of initial /l/ and /r/, and their vowel environments	90
3.4	Means of consonant formant frequencies for each speaker over 30 items	91
3.5	Spearman coefficients of correlation between formants of /l/ or /r/ and of the following vowel	98
3.6	Formant correlations as a measure of coarticulation for individual speakers	99
3.7	F ratio values of formant parameters	102
3.8	Percentage correct identification using LP spectra	105
3.9	Percentage correct identification with test LP spectra computed from 1, 2, 5 or 10 test items	107
3.10	Values of the DISTANCE measure of coarticulation	110
3.11	Rank ordering of speakers by coarticulation as assessed by formant correlations and by the DISTANCE measure	111
3.12	Spearman correlations of different measures of /l/ coarticulation	112
3.13	Spearman correlations of different measures of /r/ coarticulation	112
3.14	Percentage correct identification using coarticulation	113
4.1	Measures of characteristics of long term spectra of voice qualities	153
4.2	Means of F_1 , F_2 , F_3 for long term qualities: mp 1-16	160
4.3	Means of F_1 , F_2 , F_3 for long term qualities – speaker JL	169
4.4	Means of F_1 , F_2 , F_3 for long term qualities – speaker FN	178
4.5	Pole frequencies of vowels produced with a change in larynx height	186
4.6	Identification of voice qualities across speakers: six-value descriptor	189
4.7	Mean F_2 and F_1 , and the ratio between them, for pairs of vowels representing three vowel categories, according to speaker type	190
4.8	Mean $F_2:F_1$ ratios for vowels of three categories, for nine voice qualities realised by two speakers	190
4.9	Identification of voice qualities across speakers: three-value descriptor	191

Figures

1.1	'Voiceprint' of <i>The Phonetic Bases of Speaker Recognition</i>	19
2.1	Overview of model of sources of between-speaker differences	30
2.2	The segmental strand	39
2.3	The suprasegmental strand	47
2.4	Integration rules, phonetic representation, implementation rules and physical constraints	52
2.5	Schematic impression of part of a phonetic representation	57
2.6	Schematic representation of the use of a prestige value of a socio-linguistic variable according to class and style	64
3.1	Means (over 15 speakers) of formant frequencies of /l/ and the vowel environments	88
3.2	Means (over 13 speakers) for formant frequencies of /r/ and the vowel environments	89
3.3	Mean formant frequencies of /l/ for each speaker	92
3.4	Mean formant frequencies of /r/ for each speaker	93
3.5	Means for speaker GRP of formant frequencies of /l/ and the following vowel environment	94
3.6	Means for speaker TL of formant frequencies of /l/ and the following vowel environment	94
3.7	Means for speaker JB of formant frequencies of /l/ and the following vowel environment	95
3.8	Means for speaker NL of formant frequencies of /l/ and the following vowel environment	95
3.9	Means for speaker JRC of formant frequencies of /r/ and the following vowel environment	97
3.10	Means for speaker PVE of formant frequencies of /r/ and the following vowel environment	97
3.11	Rank ordered coefficients of correlation (r_s) for 13 speakers	100
3.12	Identification rates averaging items to form test sample	108
3.13	Rank ordering of speakers by /l/ coarticulation	109
3.14	Rank orderings of speakers by /r/ coarticulation	112
3.15	Coarticulatory implementation strategies	119
4.1	Summary of Laver's descriptive framework for long term qualities	137
4.2	Sagittal section showing stylised radial displacements of the notional centre of mass of the tongue body in settings	138
4.3a-c	Long term spectra of voice qualities (speaker JL)	144-6
4.4a-c	Long term spectra of voice qualities (speaker FN)	148-50
4.5	Long term spectra of voice qualities: amplitude ratios	152
4.6	Long term spectra of voice qualities: slope approximation	154
4.7	Means for formant frequencies for long term qualities: mp 1-16	161
4.8	Means of formant frequencies for nine long term qualities: mp 17-67 and (circled) 1-16	170
4.9	Subcategorisation of measurement points (JL)	172
4.10	Subcategorisation of measurement points (FN)	179
4.11	Tracings from x-ray pictures of the larynx and pharynx in neutral, extreme raised larynx, and extreme lowered larynx settings	183