

---

# Contents

---

<b>Acknowledgements</b>	<b>v</b>
<b>Abstract</b>	<b>vii</b>
<b>Version abrégée</b>	<b>ix</b>
<b>Contents</b>	<b>xi</b>
<b>List of Figures</b>	<b>xv</b>
<b>List of Tables</b>	<b>xx</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Recognizing individuals by their voices . . . . .	1
1.1.1 The importance of determining the identity of a speaker in forensic applications . . . . .	2
1.2 Admissibility of scientific evidence in courts . . . . .	2
1.3 Speech as evidence . . . . .	4
1.4 Forensic speaker recognition: aural-perceptual, auditory-instrumental and automatic approaches . . . . .	6
1.5 Automatic speaker recognition and the task of forensic speaker recog- nition . . . . .	7
1.6 Bayesian methodology in estimating the strength of evidence . . . . .	8
1.6.1 Mismatched recording conditions and their effect on the strength of evidence . . . . .	8
1.7 Objectives of the thesis . . . . .	9
1.8 Major contributions of the thesis . . . . .	9
1.9 Organization of the thesis . . . . .	10

<b>2</b>	<b>Forensic Automatic Speaker Recognition</b>	<b>11</b>
2.1	Bayesian and frequentist approaches to forensic recognition . . . . .	12
2.2	Automatic speaker recognition . . . . .	20
2.2.1	Features used for speaker recognition . . . . .	21
2.2.2	Modeling algorithms used for automatic speaker recognition .	25
2.2.3	Alternative techniques used in speaker recognition . . . . .	29
2.2.4	Classifying speakers according to the difficulty in recognizing them . . . . .	30
2.2.5	Using automatic speaker recognition techniques in forensic automatic speaker recognition . . . . .	31
2.2.6	Kernel-density estimation for modeling between- and within-source variability of the voice . . . . .	34
2.2.7	Double Statistical Model . . . . .	36
2.3	Influence of modern communication networks and recording conditions on automatic speaker recognition . . . . .	37
2.3.1	Other groups working on this problem of mismatched conditions in forensic automatic speaker recognition . . . . .	37
2.3.2	Compensation of channel mismatch with feature mapping and speaker model synthesis . . . . .	38
2.3.3	Score normalization techniques to handle mismatch . . . . .	40
2.3.4	Handling mismatch using auxiliary features in a Bayesian Networks approach . . . . .	43
2.4	Summary . . . . .	46
<b>3</b>	<b>Bayesian interpretation and the strength of evidence in forensic automatic speaker recognition</b>	<b>49</b>
3.1	Introduction . . . . .	49
3.2	Bayesian interpretation methodology applied to automatic speaker recognition . . . . .	50
3.2.1	Bayesian interpretation in cases with sufficient suspect reference data . . . . .	50
3.2.2	Bayesian interpretation in cases with insufficient suspect reference data . . . . .	51
3.3	Multivariate and univariate approaches to estimating the strength of evidence . . . . .	54
3.3.1	Direct method . . . . .	56
3.3.2	Scoring method . . . . .	57
3.4	Variability estimation of the strength of evidence . . . . .	62
3.5	Subsets-bootstrapping of the strength of evidence . . . . .	64

3.5.1	Verbal equivalents for the likelihood ratio intervals . . . . .	66
3.6	Significance testing of the evidence . . . . .	67
3.6.1	Possible likelihood ratio outcomes in a case . . . . .	68
3.7	Complementary measures to the strength of evidence . . . . .	69
3.8	Summary . . . . .	73
<b>4</b>	<b>The influence of mismatched recording conditions on aural and automatic estimates of the strength of evidence</b>	<b>75</b>
4.1	Mismatched recording conditions and aural and automatic speaker recognition . . . . .	76
4.2	Estimating the strength of evidence in aural and automatic speaker recognition . . . . .	78
4.3	Comparing the strength of evidence in aural and automatic methods .	82
4.4	Comparison in matched and mismatched conditions in terms of Bayesian interpretation of evidence . . . . .	83
4.5	Comparison in matched and mismatched conditions in terms of Bayesian decision theory . . . . .	85
4.6	Perceptual cues used by laypersons . . . . .	89
4.7	Relative importance of the perceptual cues . . . . .	92
4.8	Summary . . . . .	92
<b>5</b>	<b>Statistical compensation techniques for handling mismatched conditions</b>	<b>95</b>
5.1	Introduction . . . . .	95
5.2	Mismatched recording conditions and the Bayesian interpretation methodology . . . . .	96
5.3	Estimating discrimination and mismatch within a database . . . . .	98
5.4	Measuring mismatch across databases . . . . .	100
5.4.1	Mismatch and the corpus-based Bayesian interpretation methodology . . . . .	103
5.5	Compensation for mismatch across databases . . . . .	109
5.5.1	Estimation and compensation of mismatch using distribution scaling . . . . .	110
5.5.2	Methodology to construct a database to handle mismatch . . .	116
5.6	Handling a case . . . . .	122
<b>6</b>	<b>Evaluation of statistical compensation techniques</b>	<b>125</b>
6.1	Evaluation using simulated case examples from the IPSC-03 database	126

6.1.1	Evaluation of the strength of evidence in matched recording conditions . . . . .	128
6.1.2	Evaluation of the strength of evidence in mismatched recording conditions . . . . .	131
6.1.3	Performances in mismatched conditions . . . . .	134
6.1.4	Statistical compensation for mismatched conditions . . . . .	136
6.1.5	Determining the size of the scaling database (S) . . . . .	137
6.1.6	Comparison with recording suspect reference and control databases in conditions of the potential population and trace databases . . . . .	147
6.2	NFI speaker recognition evaluation through a fake case . . . . .	148
6.3	Summary . . . . .	153
<b>7</b>	<b>Discussion</b>	<b>159</b>
7.1	Handling mismatch within the Bayesian interpretation framework . .	160
7.1.1	The influence of mismatch on aural and automatic speaker recognition . . . . .	163
7.1.2	Adapting the Bayesian interpretation framework to real forensic conditions . . . . .	164
7.2	Admissibility of forensic automatic speaker recognition based on Bayesian interpretation and statistical compensation in courts . . . . .	166
7.3	Summary . . . . .	169
<b>8</b>	<b>Conclusion</b>	<b>171</b>
8.1	Handling mismatched recording conditions in the Bayesian interpretation framework . . . . .	172
8.1.1	Handling mismatch in recording conditions in corpus-based forensic speaker recognition . . . . .	172
8.1.2	Methodology for creating databases to handle mismatch . . . .	173
8.1.3	Mismatched recording conditions and their effect on aural and automatic forensic speaker recognition . . . . .	173
8.2	Applying Bayesian interpretation methodology to real forensic conditions	174
8.2.1	Scoring method and direct methods for the evaluation of the likelihood ratio . . . . .	174
8.2.2	Bayesian interpretation in cases with sufficient and insufficient suspect reference data . . . . .	174
8.2.3	Analysis of the variability of the strength of evidence . . . . .	175
8.2.4	Complementary measures to the strength of evidence . . . . .	175
8.3	Future directions . . . . .	176

---

<b>A Appendix: Description of the Polyphone IPSC-01 database</b>	<b>177</b>
<b>B Appendix: Description of the Polyphone IPSC-02 database</b>	<b>179</b>
<b>C Appendix: Description of the Polyphone IPSC-03 database</b>	<b>181</b>
<b>D Appendix: Netherlands Forensic Institute Speaker recognition evaluation</b>	<b>185</b>
D.1 NFI speaker recognition evaluation through a fake case . . . . .	185
D.1.1 Correspondence . . . . .	185
D.1.2 Description of items submitted for analysis . . . . .	186
D.1.3 Plan of Work . . . . .	186
D.1.4 Methodology . . . . .	188
D.1.5 Technical Analysis . . . . .	189
D.1.6 Results of the Technical Analysis . . . . .	191
D.1.7 Conclusion . . . . .	203
<b>Bibliography</b>	<b>215</b>
<b>Curriculum Vitae</b>	<b>217</b>