

# Table of Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Outline of This Book	4
<b>2</b>	<b>ASR: An Overview</b>	<b>7</b>
2.1	General Overview	7
2.2	Automatic Processing of Speech	10
2.3	Evaluation of ASR Systems	11
2.4	Adaptation in ASR Systems	12
<b>3</b>	<b>Pre-processing of the Speech Data</b>	<b>15</b>
3.1	A/D Conversion	15
3.2	Windowing	15
3.3	Filter Bank Analysis	18
<b>4</b>	<b>Stochastic Modelling of Speech</b>	<b>21</b>
4.1	Hidden Markov Models (HMMs)	22
4.2	Solving the Three HMM Problems	25
4.2.1	Recognition	25
4.2.2	Finding the Optimal State Sequence	26
4.2.3	Training	27
<b>5</b>	<b>Knowledge Bases of an ASR System</b>	<b>31</b>
5.1	Acoustic Models	31
5.2	Pronunciation Dictionary	32
5.3	Language Models (LMs)	35
<b>6</b>	<b>Speaker Adaptation</b>	<b>37</b>
6.1	The State of the Art in Speaker Adaptation	38
6.1.1	Feature-Based Approaches	40
6.1.2	Model-Based Approaches	41
6.2	Maximum Likelihood Linear Regression	42
6.2.1	MLLR for Small Amounts of Adaptation Data	44
6.2.2	The Weighted MLLR Approach	46
6.2.3	Implementation Issues	48

6.2.4	Experiments and Results .....	49
6.3	Summary .....	54
<b>7</b>	<b>Confidence Measures .....</b>	<b>57</b>
7.1	The State of the Art in Confidence Measures .....	58
7.1.1	Statistical Hypothesis Testing .....	59
7.1.2	Using a Set of Features .....	60
7.2	Neural Networks .....	61
7.2.1	Activation Function .....	62
7.2.2	Output Function .....	64
7.2.3	Learning in NNs .....	64
7.3	Evaluating Confidence Measures .....	66
7.4	CM Features .....	66
7.4.1	Phone-Duration Based Features .....	67
7.4.2	Additional Features .....	69
7.4.3	Combining the NN Classifier with Speaker Adaptation .....	72
7.5	Experiments and Results .....	74
7.5.1	Evaluation of the NN Classifier .....	74
7.5.2	Semi-supervised Adaptation .....	76
7.6	Summary .....	78
<b>8</b>	<b>Pronunciation Adaptation .....</b>	<b>79</b>
8.1	The State of the Art in Pronunciation Modelling .....	80
8.1.1	Rule-Based Approaches .....	82
8.1.2	Data-Driven Approaches .....	83
8.1.3	Combined Approaches .....	84
8.1.4	Miscellaneous Approaches .....	86
8.1.5	Re-training the Acoustic Models .....	86
8.2	Pronunciation Modelling of Accented and Dialect Speech ....	87
8.3	Recognising Non-native Speech .....	88
8.4	Generating Non-native Pronunciation Variants .....	94
8.4.1	Classification Trees .....	97
8.4.2	Experiments and Results .....	99
8.5	Summary .....	103
<b>9</b>	<b>Future Work .....</b>	<b>105</b>
9.1	Dynamic Selection of Pronunciation Rules .....	106
<b>10</b>	<b>Summary .....</b>	<b>109</b>
	<b>Bibliography .....</b>	<b>113</b>
	<b>Index .....</b>	<b>125</b>
	<b>Glossary .....</b>	<b>127</b>

<b>A</b>	<b>Databases and Experimental Settings</b> .....	131
A.1	The German Database .....	131
A.1.1	Pre-processing of the Speech Data .....	132
A.2	Settings for NN Training and Testing .....	132
A.3	The British English WSJ Database .....	133
A.4	ISLE Database .....	133
<b>B</b>	<b>MLLR Results</b> .....	135
<b>C</b>	<b>Phoneme Inventory</b> .....	139
C.1	German Symbol Inventory .....	140
C.2	English Symbol Inventory .....	142
C.3	Manually Derived Pronunciation Rules for the ISLE Corpus .	142

Robust Adaptation to Non-Native Accents in Automatic  
Speech Recognition

Goronzy, S.

2002, XI, 146 p., Softcover

ISBN: 978-3-540-00325-0