

Contents

Table of Contents	4
List of Figures	7
List of Tables	10
List of Acronyms	11
1 Introduction	13
1.1 Zusammenfassung	13
1.2 Summary	21
1.3 Outline	29
2 Speech Synthesis	30
2.1 Vocal Tract Models	31
2.1.1 Source-Filter Models Synthesis	31
2.1.2 Formant Synthesis	33
2.1.3 Linear Prediction Coding Synthesis	33
2.1.4 Articulatory Synthesis	34
2.2 Concatenative Models	35
2.3 Unit-Selection based Synthesis	36
2.4 Hidden Markov Models based Synthesis	38
3 Hidden Markov Models Applied to Speech Synthesis	40
3.1 Motivation for using Hidden Markov Models in Speech Synthesis	40
3.1.1 Introduction to Hidden Markov Models	41
3.1.2 History of Hidden Markov Models in Speech Synthesis	46
3.2 Main Techniques for using Hidden Markov Models in Speech Synthesis	49
3.2.1 Speech Parameter Generation Algorithm	50

3.2.2	Decision Tree Based Context Clustering	53
3.2.3	Multi-space Probability Distribution Hidden Markov Model	54
3.2.4	Mel-Cepstral Analysis Technique	55
4	European Portuguese Text-to-Speech System	58
4.1	Language Dependent Module	61
4.1.1	European Portuguese Language	61
4.1.2	Contextual Information	66
4.2	Natural Language Processing Module	67
4.2.1	Motivation for using Maximum Entropy in Natural Language Processing	71
4.2.2	Natural Language Processing Tasks for European Portuguese	77
4.3	Synthesis Module	80
4.3.1	Hidden Markov Model Training Data	84
4.3.2	The Training Process	88
4.3.3	The Synthesis Process	93
4.4	Text-to-Speech System Results and Evaluation	96
4.4.1	Introduction to the Evaluation Schemes	96
4.4.2	European Portuguese Text-to-Speech System Results	98
4.4.3	Text-to-Speech System Results for a Test Utterance	101
5	Speech Corpus for Context-Based Text-to-Speech Systems	109
5.1	Speech Corpus Methodology	110
5.2	Speech Corpus Description	112
5.2.1	The Speech Corpus by Graphemes	114
5.2.2	The Phonetic Transcription by Rules	119
5.2.3	The Phonetic Transcription with Vocalic Reduction	124
5.3	The vocalic reduction influence	127
6	Analysis for System Improvements	134
6.1	Polyglot Synthesis to resolve Foreign Words	134
6.1.1	Polyglot Synthesis	136
6.1.2	European and Brazilian Portuguese Direct Mapping	138
6.1.3	Portuguese and German Languages Direct Mapping	142
6.2	Hidden Markov Models based Speech System using Residual Signal	145
6.2.1	Obtaining the Residual Signal	146

6.2.2	Integration of the residual signal in the System	148
7	Conclusions	152
	Appendices	153
A	Speech Corpora for Natural Language Processing	154
A.1	Excerpt of the Corpus for Grapheme-to-Phoneme Conversion	154
A.2	Excerpt of the Corpus for Syllable Boundaries Detection	155
A.3	Excerpt of the Corpus for Stress Prediction	155
B	Hidden Markov Models based Speech Synthesis Data Files	156
B.1	Output File from the Natural Language Processing Module	156
B.2	List of the European Portuguese Context-dependent Feature Labels .	156
B.3	Example of a Context-dependent Feature Labels File	158
B.4	Context related questions file for European Portuguese	159
B.5	Phonemes Boundaries File Format Example	167
B.6	Fundamental Frequency File Example	168
B.7	Mel-cepstral File Example	169
B.8	Data Format File Example for the Data Training	170
B.9	Window Files for Static and Dynamic Features	171
C	Hidden Markov Models Files	172
C.1	Model Prototype Definition File	172
C.2	Examples of a Phoneme's Hidden Markov Model Files	177
C.3	Contextual Information Decision Trees	189
C.4	Language Context Information Labels File Example	193
C.5	Example of a File Resulting from the Synthesis Engine	194
D	European Portuguese Phonemes and Corresponding Graphemes	196
E	Speech Corpus for Text-to-Speech Systems	199
F	Diphones in the Speech Corpus	210
G	Table with Portuguese and German Phonemes Comparison	237
	Bibliography	239