

Contents

1	Introduction	1
1.1	Motivation	1
1.2	Prosody	3
1.2.1	Segmental versus suprasegmental aspects	3
1.2.2	Definition	3
1.2.3	Prosodic functions	5
1.3	Second language learning	5
1.3.1	Second language acquisition	6
1.3.2	Second language teaching methods	8
1.3.3	Goals of second language education	10
1.3.4	Role of learner	12
1.3.5	Segmental and suprasegmental aspects	12
1.3.6	Perception and production	14
1.3.7	Tests	14
1.4	Computer-assisted language learning	15
1.4.1	State of the art in computer-assisted language learning	16
1.4.2	Analysing the learner's productions	19
1.4.3	Feedback	22
1.4.4	Economic considerations	26
1.5	Contribution of this work	26
1.6	Structure of this work	27
2	The role of prosody in first and second language	29
2.1	Prosody in first language	29
2.1.1	Language acquisition	30
2.1.2	Word accent, phrase accents and phrase boundaries	31
2.1.3	Rhythm	37
2.1.4	Paralinguistic aspects	40
2.2	Prosody in second language acquisition	41
2.2.1	Language acquisition	41
2.2.2	Word accent, phrase accents and phrase boundaries	43
2.2.3	Rhythm	45
2.2.4	Paralinguistic aspects	46

3	Automatic assessment of prosody: state of the art	47
3.1	Analysis of prosodic events	47
3.1.1	Classification of prosodic events in a first language	47
3.1.2	Detection of learner errors	52
3.2	Assessment of rhythm in a second language	57
4	Mathematical and technological foundations	63
4.1	Acoustic short-time features	65
4.1.1	Short-time frequency analysis	65
4.1.2	Mel-frequency cepstral coefficients	66
4.1.3	Intensity	70
4.1.4	Pitch	70
4.2	Classification by parametric density estimation	74
4.3	Support vector machines	77
4.3.1	The kernel trick	78
4.3.2	Optimal separating hyperplanes	80
4.3.3	Soft margin	83
4.4	Support vector regression	89
4.5	Hidden markov models	96
4.5.1	Production probability	97
4.5.2	Most likely state sequence	101
4.5.3	Estimation	102
4.5.4	Modelling multiple events	104
5	Corpora	111
5.1	C-AuDiT	111
5.1.1	Text material	112
5.1.2	Definition of subsets	112
5.1.3	Speakers	112
5.1.4	Recordings	113
5.2	AUWL	114
5.2.1	Text material	114
5.2.2	Speakers	117
5.2.3	Recordings	117
6	Annotation	119
6.1	C-AuDiT	119
6.1.1	Detailed annotation	119
6.1.2	Panel	120
6.2	AUWL	122
6.3	Inter-rater agreement	122
7	Modelling of prosody	127
7.1	Features	128
7.1.1	Generic descriptors	130
7.1.2	Features from generic descriptors	138
7.1.3	Text dependent features	142

7.1.4	Specialized rhythm features	144
7.2	Classification and regression	145
7.2.1	Word accent position	145
7.2.2	Rhythm assessment	148
8	Experimental results	151
8.1	Segmentation	151
8.2	Word accent position	153
8.2.1	Normalization	156
8.2.2	Undefined short-time features	157
8.2.3	Context	162
8.2.4	Word level	164
8.3	Rhythm	172
8.3.1	C-AuDiT	174
8.3.2	AUWL	180
9	Discussion	187
10	Summary	191
A	Corpora details	195
A.1	C-AuDiT corpus text material	195
A.2	AUWL corpus text material	206
	List of Figures	215
	List of Tables	217
	Bibliography	219