

QUALITY OF TELEPHONE-BASED SPOKEN DIALOGUE SYSTEMS

SEBASTIAN MÖLLER
Institut für Kommunikationsakustik (IKA)
Ruhr-Universität Bochum
Germany

Contents

Preface	ix
Acknowledgments	xi
Definitions and Abbreviations	xiii
1. MOTIVATION AND INTRODUCTION	1
2. QUALITY OF HUMAN-MACHINE INTERACTION OVER THE PHONE	9
2.1 Interaction Scenarios Involving Speech Transmission	11
2.1.1 Speech Transmission Systems	13
2.1.2 Room Acoustics	17
2.1.3 Spoken Dialogue Systems	18
2.2 Interaction with Spoken Dialogue Systems	38
2.2.1 Language and Dialogue Structure in HMI	41
2.2.2 Interactive Speech Theory	43
2.2.3 Cooperativity Guidelines	47
2.3 Quality of Telephone Services Based on SDSs	51
2.3.1 QoS Taxonomy	59
2.3.2 Quality Features	63
2.3.3 Validation and Discussion	63
2.4 System Specification, Design and Evaluation	66
2.4.1 Transmission Network Planning	68
2.4.2 SDS Specification	77
2.4.3 SDS Design	79
2.4.4 System Assessment and Evaluation	86
2.5 Summary	90

3. ASSESSMENT AND EVALUATION METHODS	93
3.1 Characterization	97
3.1.1 Agent Factors	97
3.1.2 Task Factors	101
3.1.3 User Factors	102
3.1.4 Environmental Factors	103
3.1.5 Contextual factors	103
3.2 Reference	103
3.3 Data Collection	106
3.4 Speech Recognition Assessment	108
3.5 Speech and Natural Language Understanding Assessment	114
3.6 Speaker Recognition Assessment	119
3.7 Speech Output Assessment	121
3.8 SDS Assessment and Evaluation	127
3.8.1 Experimental Set-Up	128
3.8.2 Test Subjects	131
3.8.3 Experimental Task	133
3.8.4 Dialogue Analysis and Annotation	134
3.8.5 Interaction Parameters	140
3.8.6 Quality Judgments	147
3.8.7 Usability Evaluation	157
3.8.8 SDS Evaluation Examples	159
3.9 Summary	162
4. SPEECH RECOGNITION PERFORMANCE OVER THE PHONE	165
4.1 Impact of Transmission Impairments on ASR Performance	169
4.2 Transmission Channel Simulation	171
4.3 Recognizer and Test Set-Up	175
4.4 Recognition Results	179
4.4.1 Normalization	181
4.4.2 Impact of Circuit Noise	183
4.4.3 Impact of Signal-Correlated Noise	184
4.4.4 Impact of Low Bit-Rate Coding	186
4.4.5 Impact of Combined Impairments	188
4.5 E-Model Modification for ASR Performance Prediction	190
4.6 Conclusions from the Experiment	195
4.7 Summary	198

5. QUALITY OF SYNTHESIZED SPEECH OVER THE PHONE	201
5.1 Functional Testing of Synthesized Speech	202
5.2 Intelligibility and Quality in Good and Telephonic Conditions	205
5.3 Test- and System-Related Influences	206
5.4 Transmission Channel Influences	208
5.4.1 Experimental Set-Up	209
5.4.2 Results	215
5.4.3 Conclusions from the Experiment	233
5.5 Summary	234
6. QUALITY OF SPOKEN DIALOGUE SYSTEMS	237
6.1 Experimental Set-Up	240
6.1.1 The BoRIS Restaurant Information System	241
6.1.2 Speech Recognition Simulation	244
6.1.3 Measurement of Interaction Parameters	248
6.1.4 Questionnaire Design	250
6.1.5 Test Design and Data Collection	252
6.2 Analysis of Assessment and Evaluation Results	255
6.2.1 Data Pre-Analysis	256
6.2.2 Multidimensional Analysis of Quality Features	258
6.2.3 Multidimensional Analysis of Interaction Parameters	273
6.2.4 Analysis of the QoS Schematic	278
6.2.5 Impact of System Characteristics	298
6.2.6 Test-Related Issues	309
6.3 Quality Prediction Modelling	311
6.3.1 PARADISE Framework	312
6.3.2 Quality Prediction Models for Experiment 6.3 Data	321
6.3.3 Hierarchical Quality Prediction Models	337
6.3.4 Conclusion of Modelling Approaches	345
6.4 Summary	348
7. FINAL CONCLUSIONS AND OUTLOOK	351
Glossary	359
Appendices	363
A Definition of Interaction Parameters	363
B Template Sentences for Synthesis Evaluation, Exp. 5.1 and 5.2	381
C BoRIS Dialogue Structure	383
D Instructions and Scenarios	387

D.1	Instructions to Test Subjects	387
D.2	Scenarios	388
D.3	Instructions for Expert Evaluation	393
E	Questionnaires	405
E.1	Questionnaire for Experiment 6.2	405
E.2	Questionnaire for Experiment 6.3	413
	References	429
	About the Author	463
	Index	465