

SPOKEN, MULTILINGUAL AND MULTIMODAL DIALOGUE SYSTEMS

DEVELOPMENT AND ASSESSMENT

Ramón López-Cózar Delgado

*Granada University
Spain*

Masahiro Araki

*Kyoto Institute of Technology
Japan*



John Wiley & Sons, Ltd

Contents

Preface	ix
1 Introduction to Dialogue Systems	1
1.1 Human-Computer Interaction and Speech Processing	1
1.2 Spoken Dialogue Systems	2
1.2.1 <i>Technological Precedents</i>	3
1.3 Multimodal Dialogue Systems	4
1.4 Multilingual Dialogue Systems	7
1.5 Dialogue Systems Referenced in This Book	7
1.6 Area Organisation and Research Directions	11
1.7 Overview of the Book	13
1.8 Further Reading	15
2 Technologies Employed to Set Up Dialogue Systems	16
2.1 Input Interface	16
2.1.1 <i>Automatic Speech Recognition</i>	17
2.1.2 <i>Natural Language Processing</i>	22
2.1.3 <i>Face Localisation and Tracking</i>	24
2.1.4 <i>Gaze Tracking</i>	26
2.1.5 <i>Lip-reading Recognition</i>	28
2.1.6 <i>Gesture Recognition</i>	30
2.1.7 <i>Handwriting Recognition</i>	33
2.2 Multimodal Processing	34
2.2.1 <i>Multimodal Data Fusion</i>	34
2.2.2 <i>Multimodal Data Storage</i>	36
2.2.3 <i>Dialogue Management</i>	41
2.2.4 <i>Task Module</i>	41
2.2.5 <i>Database Module</i>	42
2.2.6 <i>Response Generation</i>	43
2.3 Output Interface	44
2.3.1 <i>Graphic Generation</i>	44
2.3.2 <i>Natural Language Generation</i>	47
2.3.3 <i>Speech Synthesis</i>	48
2.3.4 <i>Sound Generation</i>	51
2.3.5 <i>Tactile/Haptic Generation</i>	51
2.4 Summary	51
2.5 Further Reading	53

3	Multimodal Dialogue Systems	54
3.1	Benefits of Multimodal Interaction	54
3.1.1	<i>In Terms of System Input</i>	54
3.1.2	<i>In Terms of System Processing</i>	56
3.1.3	<i>In Terms of System Output</i>	58
3.2	Development of Multimodal Dialogue Systems	59
3.2.1	<i>Development Techniques</i>	59
3.2.2	<i>Data Fusion</i>	63
3.2.3	<i>Architectures of Multimodal Systems</i>	67
3.2.4	<i>Animated Agents</i>	70
3.2.5	<i>Research Trends</i>	79
3.3	Summary	84
3.4	Further Reading	85
4	Multilingual Dialogue Systems	86
4.1	Implications of Multilinguality in the Architecture of Dialogue Systems	86
4.1.1	<i>Consideration of Alternatives in Multilingual Dialogue Systems</i>	86
4.1.2	<i>Interlingua Approach</i>	91
4.1.3	<i>Semantic Frame Conversion Approach</i>	92
4.1.4	<i>Dialogue-Control Centred Approach</i>	94
4.2	Multilingual Dialogue Systems Based on Interlingua	95
4.2.1	<i>MIT Voyager System</i>	95
4.2.2	<i>MIT Jupiter System</i>	98
4.2.3	<i>KIT System</i>	100
4.3	Multilingual Dialogue Systems Based on Web Applications	107
4.3.1	<i>Requirements for Practical Multilingual Dialogue Systems</i>	108
4.3.2	<i>Dialogue Systems Based on Web Applications</i>	108
4.3.3	<i>Multilingual Dialogue Systems Based on the MVC Framework</i>	111
4.3.4	<i>Implementation of Multilingual Voice Portals</i>	114
4.4	Summary	117
4.5	Further Reading	117
5	Dialogue Annotation, Modelling and Management	118
5.1	Dialogue Annotation	118
5.1.1	<i>Annotation of Spoken Dialogue Corpora</i>	118
5.1.2	<i>Annotation of Multimodal Dialogue Corpora</i>	121
5.2	Dialogue Modelling	124
5.2.1	<i>State-Transition Networks</i>	125
5.2.2	<i>Plans</i>	126
5.3	Dialogue Management	127
5.3.1	<i>Interaction Strategies</i>	127
5.3.2	<i>Confirmation Strategies</i>	128
5.4	Implications of Multimodality in the Dialogue Management	131
5.4.1	<i>Interaction Complexity</i>	131
5.4.2	<i>Confirmations</i>	133
5.4.3	<i>Social and Emotional Dialogue</i>	134
5.4.4	<i>Contextual Information</i>	135
5.4.5	<i>User References</i>	137
5.4.6	<i>Response Generation</i>	140

5.5 Implications of Multilinguality in the Dialogue Management	141
5.5.1 <i>Reference Resolution in Multilingual Dialogue Systems</i>	141
5.5.2 <i>Ambiguity of Speech Acts in Multilingual Dialogue Systems</i>	142
5.5.3 <i>Differences in the Interactive Behaviour of Multilingual Dialogue Systems</i>	143
5.6 Implications of Task Independency in the Dialogue Management	144
5.6.1 <i>Dialogue Task Classification</i>	144
5.6.2 <i>Task Modification in Each Task Class</i>	146
5.7 Summary	149
5.8 Further Reading	150
6 Development Tools	151
6.1 Tools for Spoken and Multilingual Dialogue Systems	151
6.1.1 <i>Tools to Develop System Modules</i>	151
6.1.2 <i>Web-Oriented Standards and Tools for Spoken Dialogue Systems</i>	159
6.1.3 <i>Internet Portals</i>	170
6.2 Standards and Tools for Multimodal Dialogue Systems	176
6.2.1 <i>Web-Oriented Multimodal Dialogue</i>	176
6.2.2 <i>Face and Body Animation</i>	179
6.2.3 <i>System Development Tools</i>	181
6.2.4 <i>Multimodal Annotation Tools</i>	185
6.3 Summary	187
6.4 Further Reading	187
7 Assessment	189
7.1 Overview of Evaluation Techniques	189
7.1.1 <i>Classification of Evaluation Techniques</i>	190
7.2 Evaluation of Spoken and Multilingual Dialogue Systems	192
7.2.1 <i>Subsystem-Level Evaluation</i>	192
7.2.2 <i>End-to-End Evaluation</i>	196
7.2.3 <i>Dialogue Processing Evaluation</i>	197
7.2.4 <i>System-to-System Automatic Evaluation</i>	199
7.3 Evaluation of Multimodal Dialogue Systems	202
7.3.1 <i>System-Level Evaluation</i>	203
7.3.2 <i>Subsystem-Level Evaluation</i>	208
7.3.3 <i>Evaluation of Multimodal Data Fusion</i>	210
7.3.4 <i>Evaluation of Animated Agents</i>	212
7.4 Summary	217
7.5 Further Reading	218
Appendix A Basic Tutorial on VoiceXML	219
Appendix B Multimodal Databases	229
Appendix C Coding Schemes for Multimodal Resources	233
Appendix D URLs of Interest	235
Appendix E List of Abbreviations	237
References	239
Index	253