



# Voice Processing

---

Gordon E. Pelton

Uyless Black,  
Consulting Editor

McGraw-Hill Series on Computer Communications

# Contents

Preface	xi
Acknowledgments	xv
<b>Chapter 1. Introduction</b>	
Speech Technologies	
Voice Processing Applications	
Nontelephone Applications	
Call Processing	
Voice Messaging	
Interactive Voice Response (IVR)	
Application Acquisition	
<b>Chapter 2. History of Talking Automata</b>	
Speaking Idols of the Ancients	13
Talking Heads in the Middle Ages	14
Some Contrivances of the Seventeenth and Eighteenth Centuries	15
Professor Kratzenstein's Prize	15
The Speaking Machines of Wolfgang von Kempelen	16
Two Talking Heads of the Abbe' Mical	21
Further Efforts in the Nineteenth Century	22
The Automaton Speaking Figure of Professor Joseph Faber	22
The Contributions of Alexander Graham Bell	23
Developments in the Early Twentieth Century	26
Pedro, the VODER	27
Milestones in the World of Telephone Communications	32
The Digital Computer and Voice Processing	32

<b>Chapter 3. Speech Technology</b>	<b>33</b>
Production of Speech Waveforms	34
Digitization of Analog Waveforms	37
Sampling Analog Waveforms	38
Quantizing Analog Waveforms	40
Analysis/Synthesis in the Frequency Domain	47
Spectrum Analysis of Speech Waveforms	48
Subband Coding—SBC	54
Analysis/Synthesis in the Time Domain	55
Differential Coding	56
Evaluation of Coding Schemes	66
Text-to-Speech	67
Linguistic Issues	68
The Conversion Process	72
<b>Chapter 4. Speech Recognition</b>	<b>83</b>
Variations in Speech	84
Classification of Systems	85
Speaker Dependence	86
Vocabulary Size	87
Telephone Speech	88
Speech Recognition Technology	88
Hidden Markov Models—HMM	90
Choice of Speech Units	93
Modeling Speech Units with Hidden Markov Models	96
Modeling Grammar with Hidden Markov Models	97
Acoustic/Phonetic Example Using Hidden Markov Model	99
Measures of Performance	102
<b>Chapter 5. Telecommunications</b>	<b>103</b>
Historical Perspective	103
Telecommunications Networks	106
Switched Network	108
Central Office (CO)	109
Remote Switch	110
Toll Office	111
Tandem Office	111
Toll Network Switching Hierarchy	113
Public Versus Private Telephone Networks	115
Analog Versus Digital Networks	117
Integrated Services Digital Network (ISDN)	118
Transmission	119
Transmission Media	120
Modulation	124
Multiplexing	132
Carrier Systems	136
Cellular Systems	140
Switching	141
Manual Switching	141

Electromechanical Switching	143
Digital Switching	151
Private Branch Exchange (PBX)	159
Signaling	162
A Typical Call	164
Subscriber Loop Signaling	165
Interoffice Signaling	170
<b>Chapter 6. Requirements and Features</b>	<b>173</b>
Requirements	174
Basic Capabilities	174
Miscellaneous Capabilities	177
External Interfaces	184
Optional Installed Hardware	184
Design Features	186
Screen and Keyboard Versus Telephone	186
Systemwide Features	188
Planning for Data Input	189
Miscellaneous Ideas	191
Script Design	191
Scenarios	193
Prompt and Menu Design	193
<b>Chapter 7. Hardware</b>	<b>197</b>
Microcircuits	197
Telecommunications	198
Signal Processing	198
Digital Signal Processors (DSP)	198
Portable, Special-Purpose Speech DSP	200
Personal Computer Add-in Boards	201
Telephone/Voice Boards	202
Telephony Interfaces	209
Speech Recognition	212
Facsimile (FAX)	215
Text-to-Speech	216
Peripheral Text-to-Speech Products	219
System-Level Products	220
<b>Chapter 8. Development Software</b>	<b>221</b>
Device Drivers	222
Multiline Requirements	223
Special Features	225
User Interface	227
Function Libraries for Higher-Level Languages	229
Multitasking Systems	229
Requirements	229
Voice Processing Systems	230
Embedded Multitasking	231

Multitasking Executives	233
Development Toolkits	238
Features	239
Components	240
Application Generators	246
Overview	246
Features	252
Software Architecture	253
Vocabulary Production	275
Application Documentation	276
Creating an APGEN from a Toolkit	276
Speech Editors	277
Hardware Requirements	278
Speech File Management	278
Recording	280
Visual Editing	281
<b>Chapter 9. Vocabulary Production</b>	<b>289</b>
Permanent Vocabulary Production	290
The Recording Facility	292
Some Principles of Audio	292
The Professional Studio	297
Recording Studio at Home or Office	297
Hardware for Vocabulary Production	298
Microphone Characteristics	299
Microphone Types	305
Recording Equipment	308
Digitization and Editing Devices	309
The Announcer	312
Candidates	313
Auditions	313
Candidate Evaluation	314
The Recording Session	314
Preparation	314
Script Design	315
Microphone Techniques	315
Direction of a Recording Session	316
Prototype Versus Permanent Vocabulary	317
Vocabulary Editing	317
The Process	318
Marginal Utterances	318
Indexed Files	319
Final Review	319
Conclusion	319
<b>Appendix A. MAX, The Check-in System</b>	<b>321</b>
Telephone User's Guide	321
All employees	321
Supervisors	323

<b>Appendix B. Amy Scheduling System</b>	<b>325</b>
<b>Appendix C. Sample Scenario for an Auction-Results Telephone System</b>	<b>331</b>
<b>A Possible Scenario</b>	<b>331</b>
<b>Caller has not submitted an absentee bid</b>	<b>332</b>
<b>Caller has submitted an absentee bid</b>	<b>332</b>
<b>Appendix D. Top-Level Dialogue for Guard Check-in System</b>	<b>337</b>
<b>Appendix E. C-Language Program Using a Voice Processing Development Toolkit</b>	<b>341</b>
<b>Appendix F. Script Program in Assembler-Like Language</b>	<b>353</b>
<b>Appendix G. Sample Script Language Program</b>	<b>365</b>
<b>Bibliography</b>	<b>375</b>
<b>Index</b>	<b>381</b>