Contents

Preface ix

Chapter 1. Sound 1

1.1. Acoustic phonetics 1
1.2. Sound waves 2
1.3. In search of pure sound 3
1.4. Amplitude, frequency, duration and phase 6
1.4.1. Amplitude 6
1.4.2. Frequency 7
1.4.3. Duration 8
1.4.4. Phase 8
1.5. Units of pure sound 9
1.6. Amplitude and intensity 10
1.7. Bels and decibels 12
1.8. Audibility threshold and pain threshold 13
1.9. Intensity and distance from the sound source 14
1.10. Pure sound and musical sound: the scale in Western music 15
1.11. Audiometry 16
1.12. Masking effect 17
1.13. Pure untraceable sound 18
1.14. Pure sound, complex sound 19

Chapter 2. Sound Conservation 25

2.1. Phonautograph 25
2.2. Kymograph 27

2.3. Recording chain 31
2.3.1. Distortion of magnetic tape recordings 32
2.3.2. Digital recording 34
2.4. Microphones and sound recording 35
2.5. Recording locations 36
2.6. Monitoring 37
2.7. Binary format and Nyquist–Shannon frequency 38
2.7.1. Amplitude conversion 38
2.7.2. Sampling frequency 39
2.8. Choice of recording format 40
2.8.1. Which sampling frequency should be chosen? 40
2.8.2. Which coding format should be chosen? 41
2.8.3. Recording capacity 41
2.9. MP3, WMA and other encodings 42

Chapter 3. Harmonic Analysis 45

3.1. Harmonic spectral analysis 45
3.2. Fourier series and Fourier transform 53
3.3. Fast Fourier transform 54
3.4. Sound snapshots 54
3.5. Time windows 55
3.6. Common windows 57
3.7. Filters 59
3.8. Wavelet analysis 60
3.8.1. Wavelets and Fourier analysis 60
3.8.2. Choice of the number of cycles 61

Chapter 4. The Production of Speech Sounds 65

4.1. Phonation modes 65
4.2. Vibration of the vocal folds 68
4.3. Jitter and shimmer 72
4.4. Friction noises 72
4.5. Explosion noises 73
4.6. Nasals 74
4.7. Mixed modes 74
4.8. Whisper 74
4.9. Source-filter model 75

Chapter 5. Source-filter Model Analysis 77

5.1. Prony's method – LPC 77
5.1.1. Zeros and poles 79
5.2. Which LPC settings should be chosen? 81
5.2.1. Window duration? 81
5.2.2. What order for LPC? 82
5.3. Linear prediction and Prony's method: nasals 82
5.4. Synthesis and coding by linear prediction 83

Chapter 6. Spectrograms 85

6.1. Production of spectrograms 85
6.2. Segmentation 88
6.2.1. Segmentation: an awkward problem (phones, phonemes, syllables, stress groups) 88
6.2.2. Segmentation by listeners 89
6.2.3. Traditional manual (visual) segmentation 90
6.2.4. Phonetic transcription 92
6.2.5. Silences and pauses. 92
6.2.6. Fricatives 93
6.2.7. Occlusives, stop consonants. 95
6.2.8. Vowels 97
6.2.9. Nasals. 99
6.2.10. The R 100
6.2.11. What is the purpose of segmentation? 101
6.2.12. Assessment of segmentation. 101
6.2.13. Automatic computer segmentation 101
6.2.14. On-the-fly segmentation 103
6.2.15. Segmentation by alignment with synthetic speech. 104
6.2.16. Spectrogram reading using phonetic analysis software 106
6.3. How are the frequencies of formants measured? 106
6.4. Settings: recording. 112

Chapter 7. Fundamental Frequency and Intensity 117

7.1. Laryngeal cycle repetition 117
7.2. The fundamental frequency: a quasi-frequency. 119
7.3. Laryngeal frequency and fundamental frequency 120
7.4. Temporal methods. 123
7.4.1. Filtering 126
7.4.2. Autocorrelation 127
7.4.3. AMDF 129

7.5. Frequency (spectral) methods 129
7.5.1. Cepstrum 130
7.5.2. Spectral comb 132
7.5.3. Spectral brush 133
7.5.4. SWIPE 133
7.5.5. Measuring errors of F0 134
7.6. Smoothing 135
7.7. Choosing a method to measure F0 138
7.8. Creaky voice 139
7.9. Intensity measurement 142
7.10. Prosodic annotation 143
7.10.1. Composition of accent phrases 144
7.10.2. Annotation of stressed syllables, mission impossible? 145
7.10.3. Framing the annotation of stressed syllables 146
7.10.4. Pitch accent 147
7.10.5. Tonal targets and pitch contours 148
7.11. Prosodic morphing 148
7.11.1. Change in intensity 149
7.11.2. Change in duration by the Psola method 149
7.11.3. Slowdown/acceleration 150
7.11.4. F0 modification 152
7.11.5. Modification of F0 and duration by phase vocoder 153

Chapter 8. Articulatory Models 157

8.1. History 157
8.2. Single-tube model 160
8.3. Two-tube model 162
8.4. Three-tube model 167
8.5. N-tube model 172

Appendix 173

References 177

Index 181