T-Labs Series in Telecommunication Services

Series Editors

Sebastian Möller, TU Berlin and Telekom Innovation Laboratories, Berlin, Germany
Axel Küpper, TU Berlin and Telekom Innovation Laboratories, Berlin, Germany
Alexander Raake, TU Berlin and Telekom Innovation Laboratories, Berlin, Germany

For further volumes:
http://www.springer.com/series/10013
Dimension-based Quality Modeling of Transmitted Speech
Preface

In the present book, speech transmission quality is modeled on the basis of perceptual dimensions that are relevant for today’s public-switched and packet-based telecommunication systems. The complete transmission path from the mouth of the speaker to the ear of the listener is regarded, and both narrowband (300–3400 Hz) as well as wideband (50–7000 Hz) speech transmission are taken into account. A new analytical assessment method is developed that allows the dimensions to be rated by non-expert listeners, and a new parametric model for the quality estimation of transmitted speech based on the perceptual dimensions is derived.

This book was created within the scope of my dissertation at the Institute of Communication Acoustics (IKA) at Ruhr-University Bochum, and to the largest extent at the Quality and Usability Lab at TU Berlin, which in turn is part of Telekom Innovation Laboratories (T-Labs). During the last years, a large number of individuals supported my scientific activities at both professional and personal levels. First of all, I owe my deepest gratitude to my colleague and supervisor Prof. Sebastian Möller for enabling this work, for his advice, for the scientific freedom he provided, for many fruitful discussions, and for the ongoing motivation, his patience, and his trust also in difficult times. I am truly indebted to Prof. Ulrich Heute for his interest in my work, his ideas, numerous inspiring discussions, and for his willingness to co-supervise and examine the present work. I would also like to thank my colleague Prof. Alexander Raake for the longstanding company through different stages of my scientific career, his ideas and the motivation, and for numerous intensive and in-depth discussions. I would not have met these excellent teachers if Prof. Jens Blauert had not introduced me to the fascinating topics of technical acoustics and communication acoustics in his Bochum lectures when I was an undergraduate, which I greatly appreciate.

I am particularly thankful to Kirstin Scholz for the productive and exciting collaboration over several years, for sharing her thoughts in countless valuable discussions, and for the review of large parts of an earlier version of this book. The close collaboration with Nicolas Côté was always a pleasure for which I would like to express him my gratitude. I am also very grateful to Lu Huo and Błażej Lewcio for contributing their expertise in many joint research efforts.
I would like to acknowledge the very friendly, diverse, and supportive atmosphere at T-Labs and I am obliged to many of my colleagues who supported and advised me in different concerns, in particular Ina Wechsung, Robert Schleicher, Sascha Spors, Jens Ahrens, Tim Polzehl, Benjamin Weiss, Benjamin Belmudez, Jan-Niklas Antons, Christine Kühnel, Marie-Neige Garcia, Anja Naumann, Matthias Geier, Juan-Pablo Ramírez, Klaus-Peter Engelbrecht, Pablo Vidalés, and Niklas Kirschnich. Moreover, I would like to thank Irene Hube-Achter, Tobias Hirsch, Klaus-Jürgen Buß, and their great teams for the administrative support and for providing more than the obligatory resources. I am also very grateful to Falk Schiffner and Ulrike Stiefelhagen for the organization and realization of many auditory experiments. Besides these individuals, I would also like to thank all the “new” colleagues for their impressive support during my doctoral defence.

It is a great pleasure to thank Jens Berger, Hans-Wilhelm Gierlich, Vincent Barriac, and other colleagues from Study Group 12 of the International Telecommunication Union (ITU-T) for sharing their expertise and for being so open-minded to newcomers.

For their financial support, I am grateful to the Deutsche Forschungsgemeinschaft (DFG).

My special thanks go to my family and friends (to which many of the above-mentioned colleagues belong) for their empathy, their moral support, and for the joyful moments during the last years. Finally, and most important, I would sincerely like to thank Joanna. This work would not have been possible without all her support, her patience, and her ongoing encouragement.

Berlin, September 2012
Marcel Wältermann
Contents

1 Introduction ........................................ 1

2 A Dimension-Based Approach to Mouth-to-Ear Speech Transmission Quality ........................................ 5
  2.1 General Research Scenario ............................. 5
  2.2 Speech Transmission in Telecommunication ............ 9
      2.2.1 Introduction .................................... 9
      2.2.2 Mouth to Channel ............................... 10
      2.2.3 Channel .......................................... 18
      2.2.4 Channel to Ear ................................. 19
  2.3 Perception of Transmitted Speech ...................... 22
      2.3.1 Introduction ..................................... 22
      2.3.2 Schematic of a Listener ......................... 23
      2.3.3 Perceptual Features and Dimensions .............. 25
      2.3.4 Integral Quality, Quality Features and Dimensions, and Quality Elements ............................ 26
      2.3.5 QoS and QoE Terminology ....................... 28
  2.4 Auditory Quality Measurement ......................... 29
      2.4.1 Listener as a Measuring Organ .................. 29
      2.4.2 Scaling Functions ............................... 30
      2.4.3 Psychometric Methods ........................... 31
      2.4.4 Personal and External Modifying Factors and Some Countermeasures .......................... 38
      2.4.5 Scale Transformation ............................ 42
      2.4.6 Towards a Universal Continuum for Perceptual Value ........................................ 45
2.5 Dimension-Based Quality Models ........................................... 46
  2.5.1 Principle ................................................................. 46
  2.5.2 Vector Model and Ideal-Point Model ................................. 47
  2.5.3 Combination of Impairments ........................................ 50
2.6 Instrumental Quality Measurement ......................................... 51
  2.6.1 Introduction ............................................................. 51
  2.6.2 Signal-Based Instrumental Models ................................ 54
  2.6.3 The E-Model, a Parametric Instrumental Model ................. 56
2.7 Research Topics Covered in this Book .................................. 60

3 Quality Feature Space of Transmitted Speech .............................. 63
  3.1 Introduction ............................................................... 63
  3.2 Experimental Paradigms .................................................. 64
    3.2.1 Pairwise Similarity and MDS .................................... 65
    3.2.2 Semantic Differential and PCA ................................... 66
    3.2.3 Three-way Models .................................................. 67
  3.3 Literature Review: Speech-Quality Features and Dimensions ............. 68
  3.4 Experimental Set-Up ..................................................... 73
    3.4.1 Speech Samples ..................................................... 73
    3.4.2 Test Room and Participants ..................................... 76
  3.5 Determination of SD Attributes ...................................... 77
  3.6 Multidimensional Analyses ............................................. 78
    3.6.1 Speaker/Sentence- and Subject-dependency of the Data ............ 78
    3.6.2 Resulting Perceptual Dimensions ................................ 80
  3.7 Relevance for Quality .................................................. 88
  3.8 Conclusions ............................................................... 90
  3.9 Considerations on a “Loudness Feature” ................................ 91
  3.10 Global Dimensions Versus Local Dimensions ............................ 93

4 Direct Scaling of Speech Quality Dimensions ............................... 95
  4.1 Introduction ................................................................ 95
  4.2 Dimension Rating Scales ................................................ 96
  4.3 Test Procedure .......................................................... 97
    4.3.1 General ................................................................. 97
    4.3.2 Dimension Assessment ............................................. 98
    4.3.3 Number of Test Stimuli and Overall Test Duration ............... 100
    4.3.4 Details on the Experiment Organization .......................... 100
  4.4 Application ............................................................... 102
    4.4.1 Speech Samples ..................................................... 102
    4.4.2 Test Room and Participants ..................................... 104
4.5 Results and Analysis .................................................. 104
4.5.1 General Characteristics of the Data .......................... 104
4.5.2 Speaker/Sentence- and Subject-Dependency of the Data .......................... 106
4.5.3 Relevant Effects .................................................. 106
4.5.4 Example Quality and Dimension Scores ....................... 108
4.6 Conclusions ......................................................... 111

5 Instrumental Dimension-Based Speech Quality Modeling ........ 115
5.1 Introduction ......................................................... 115
5.2 Dimension-Based Quality Model .................................. 116
5.2.1 Total Impairment ................................................. 116
5.2.2 Dimension Impairment Factors ................................. 118
5.2.3 Distance Model .................................................. 120
5.3 Instrumental Dimension Models ................................... 125
5.3.1 Introduction ..................................................... 125
5.3.2 Estimation of “Discontinuity” ................................. 125
5.3.3 Estimation of “Noisiness” ...................................... 133
5.3.4 Estimation of “Coloration” .................................... 145
5.4 Dimension-Based Estimation of Integral Quality: The DNC-Model ........................................ 150
5.4.1 Introduction ..................................................... 150
5.4.2 DNC-Model ..................................................... 151
5.4.3 Evaluation ....................................................... 152
5.5 Signal-Based Instrumental Quality Models Based on Dimensions ........................................ 158
5.6 Conclusions ......................................................... 160

6 Conclusions and Future Work ........................................ 163

Appendix A: Logistic and Log-Logistic Functions .................... 169
Appendix B: E-Model Algorithms ...................................... 173
Appendix C: Experiments ................................................ 181

References ........................................................... 195
<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABE</td>
<td>Artificial Bandwidth Extension</td>
</tr>
<tr>
<td>ACR</td>
<td>Absolute Category Rating</td>
</tr>
<tr>
<td>ADPCM</td>
<td>Adaptive Differential Pulse Code Modulation</td>
</tr>
<tr>
<td>AEC</td>
<td>Acoustic Echo Cancelation</td>
</tr>
<tr>
<td>AMR</td>
<td>Adaptive Multi-rate</td>
</tr>
<tr>
<td>ANOVA</td>
<td>Analysis of Variance</td>
</tr>
<tr>
<td>BeTR</td>
<td>Bellcore Transmission Rating Model</td>
</tr>
<tr>
<td>CANDECOMP</td>
<td>Canonical Decomposition</td>
</tr>
<tr>
<td>CELP</td>
<td>Code-excited Linear Prediction</td>
</tr>
<tr>
<td>CI</td>
<td>Confidence Interval</td>
</tr>
<tr>
<td>DAM</td>
<td>Diagnostic Acceptability Measure</td>
</tr>
<tr>
<td>DECT</td>
<td>Digital Enhanced Cordless Telecommunications</td>
</tr>
<tr>
<td>DIAL</td>
<td>Diagnostic Instrumental Assessment of Listening-quality</td>
</tr>
<tr>
<td>DIF</td>
<td>Dimension Impairment Factor</td>
</tr>
<tr>
<td>DNC</td>
<td>Discontinuity-Noisiness-Coloration</td>
</tr>
<tr>
<td>DSL</td>
<td>Digital Subscriber Line</td>
</tr>
<tr>
<td>DTX</td>
<td>Discontinuous Transmission</td>
</tr>
<tr>
<td>ERP</td>
<td>Ear Reference Point</td>
</tr>
<tr>
<td>ETSI</td>
<td>European Telecommunications Standards Institute</td>
</tr>
<tr>
<td>FB</td>
<td>Fullband (20–20000 Hz)</td>
</tr>
<tr>
<td>FEC</td>
<td>Forward Error Correction</td>
</tr>
<tr>
<td>FIR</td>
<td>Finite Impulse Response</td>
</tr>
<tr>
<td>GSM</td>
<td>Global System for Mobile Communications</td>
</tr>
<tr>
<td>HATS</td>
<td>Head and Torso Simulator</td>
</tr>
<tr>
<td>HFT</td>
<td>Hands-free Terminal</td>
</tr>
<tr>
<td>IDIOSCAL</td>
<td>Individual Differences in Orientation Scaling</td>
</tr>
<tr>
<td>INDSCAL</td>
<td>Individual Differences Scaling</td>
</tr>
<tr>
<td>IRS</td>
<td>Intermediate Reference System</td>
</tr>
<tr>
<td>ISDN</td>
<td>Integrated Services Digital Network</td>
</tr>
<tr>
<td>LBR</td>
<td>Low-bitrate Redundancy</td>
</tr>
<tr>
<td>LP</td>
<td>Linear Prediction</td>
</tr>
</tbody>
</table>
LPC  Linear Predictive Coding
LTI  Linear Time-invariant
LTP  Long-term Predictor
MANOVA  Multivariate Analysis of Variance
MDS  Multidimensional Scaling
MNRU  Modulated Noise Reference Unit
MOS  Mean Opinion Score
MRP  Mouth Reference Point
NB  Narrowband (300–3400 Hz)
NR  Noise Reduction
OLR  Overall Loudness Rating
OSI  Open Systems Interconnection
PARAFAC  Parallel Factor Analysis
PC  Principal Component
PCA  Principal Component Analysis
PCM  Pulse Code Modulation
PESQ  Perceptual Evaluation of Speech Quality
PLC  Packet-loss Concealment
POLQA  Perceptual Objective Listening Quality Assessment
PS  Pairwise Similarity
PSTN  Public Switched Telephone Network
QoE  Quality of Experience
QoS  Quality of Service
RELP  Residual-excited Linear Prediction
RLR  Receive Loudness Rating
RTCP  RTP Control Protocol
RTP  Real-time Transport Protocol
SD  Semantic Differential
SDP  Session Description Protocol
SIP  Session Initiation Protocol
SNR  Signal-to-noise Ratio
SLR  Send Loudness Rating
SVD  Singular Value Decomposition
SWB  Super-wideband (50–14000 Hz)
TCP  Transmission Control Protocol
TOSQA  Telecommunication Objective Speech-Quality Assessment
UDP  User Datagram Protocol
UMTS  Universal Mobile Telecommunications System
VAD  Voice Activity Detection
VoIP  Voice over Internet Protocol
WB  Wideband (50–7000 Hz)