

# **T-Labs Series in Telecommunication Services**

## **Series editors**

Sebastian Möller, Berlin, Germany

Axel Küpper, Berlin, Germany

Alexander Raake, Berlin, Germany

More information about this series at <http://www.springer.com/series/10013>

Jan-Niklas Antons

# Neural Correlates of Quality Perception for Complex Speech Signals

 Springer

Jan-Niklas Antons  
Quality and Usability Lab  
Technische Universität Berlin  
Berlin  
Germany

Zugl.: Berlin, Technische Universität, Diss., 2014

ISSN 2192-2810 ISSN 2192-2829 (electronic)  
T-Labs Series in Telecommunication Services  
ISBN 978-3-319-15520-3 ISBN 978-3-319-15521-0 (eBook)  
DOI 10.1007/978-3-319-15521-0

Library of Congress Control Number: 2015930729

Springer Cham Heidelberg New York Dordrecht London  
© Springer International Publishing Switzerland 2015

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made.

Printed on acid-free paper

Springer International Publishing AG Switzerland is part of Springer Science+Business Media  
([www.springer.com](http://www.springer.com))

# Preface

This book presents the research of the author on the neural correlates of quality perception for complex speech signals. Two different disciplines will be interconnected here, namely neuroscience and Quality of Experience research, which do not seem to be frequently used in combination for research on speech quality perception. In the five experiments conducted here, standard clinical methods in neurophysiology on the one hand, and on the other hand, methods used in fields of research concerned with speech quality perception, will be applied. Using this combination, it will be shown that speech stimuli with different lengths (phonemes, words, sentences and audiobooks) and different quality impairments (signal-correlated noise, reduced bit rate of a speech codec and reverberation) are accompanied by physiological reactions related to quality variations, e.g. a positive peak in an event-related potential. Furthermore, it will be shown that—in most cases—quality impairment intensity has an impact on the strength of the intensity of physiological reactions (components of event-related potentials in Chaps. 2–4, or alpha frequency band power in Chaps. 5, and 6). This book consists of the following contributions: Implementation of a test set-up combining neurophysiological and subjective quality assessment methods for speech quality perception testing (Chaps. 2–6). The proof that this test set-up successfully functions with short speech stimuli (phonemes) and generic quality impairment, i.e. signal-correlated noise (Chap. 2). A successful application of this test method to longer speech stimuli (words) with a more realistic quality impairment, i.e. reduced bit rate of a speech codec (Chap. 3). The proof that this technique successfully functions in respect to stimuli with lengths for standard quality testing (sentences) and an environment-related quality impairment, i.e. reverberation (Chap. 4). An investigation of the impact of a speech compression algorithm with reduced bit rate on the cognitive state of listeners for speech stimuli of long duration (audiobooks) in constant (Chap. 5) and varying quality conditions (Chap. 6).

Berlin, December 2014

Jan-Niklas Antons

# Acknowledgments

This work would not have been possible without the help of numerous supporters. Thank you to everyone who supported me during this work. A special mention belongs to the following institutions and persons:

- I am grateful to the Technische Universität Berlin, the Telekom Innovation Laboratories, and the Bernstein Focus: Neurotechnologie—Berlin which provided the foundation for all my work. I am especially thankful for the effortful work of Dr. Heinrich Arnold, Prof. Dr. Klaus-Robert Müller, Prof. Dr. Benjamin Blankertz and their teams: thank you for making this work possible.
- I am grateful to the Quality and Usability Lab, the group Assessment of IP-based Applications, and all the colleagues who supported me over the years. You let me experience how it feels to work in a great team.
- I am grateful to my student workers Ahmad Abbas and Steffen Zander who supported me brilliantly over the years.
- I am grateful to Sebastian Arndt, Dr.-Ing. Benjamin Belmudez, Dr.-Ing. Marcel Wältermann, Prof. Dr.-Ing. Alexander Raake, Dr.-Ing. Tim Polzehl, Dr. Benjamin Weiss, Dr.-Ing. Marie-Neige Garcia, and Dr.-Ing. Jens Ahrens with whom I spent much time discussing and working in one on the most interesting research areas.
- I am grateful to Irene Hube-Achter, Yasmin Hillebrenner, and Tobias Hirsch whom organized the Quality and Usability Lab so brilliant over all the years and helped me in every situation with a well-suited solution.
- I am grateful to Dr. Robert Schleicher who supported me from the first day of my postgraduate time in uncountable manners. Thanks for the good time and your true understanding in so many situations.
- I am grateful to the reviewers of my doctoral thesis: Prof. Dr. med. Gabriel Curio and Prof. Tiago H. Falk, Ph.D., for their scientific and thesis-related support over the last years.
- I am grateful to my supervisor Sebastian Möller who supported me not only in every research-related question but also showed me how to organize (business-) life.

- I am grateful to Viktoria Voigt for supporting me in all possible ways: scientific, business and private matters. Thank you for your love and for understanding me. Thanks for pointing me in the right direction when I am too blind to find it on my own.
- I am grateful to my parents for the years of support during my entire life. You are the foundation of my entire life, thanks for being there for me and understanding me. Thank you for teaching me strength and strong will.

# Contents

<b>1</b>	<b>Introduction</b>	1
1.1	Quality and Quality of Experience	4
1.2	Quality Assessment Methods for Speech Stimuli	7
1.3	Electrophysiology and Electroencephalogram	10
1.3.1	EEG Frequency Band Power	11
1.3.2	Working Memory, Vigilance, and Cognitive State	13
1.3.3	EEG Experiment	14
1.3.4	Event-Related Potentials	16
1.4	Outline and Objective of this Work	24
<b>2</b>	<b>ERPs and Quality Ratings Evoked by Phoneme Stimuli Under Varying SNR Conditions</b>	27
2.1	Introduction	28
2.2	Methods	29
2.2.1	Participants	29
2.2.2	Material	29
2.2.3	Experimental Design and Procedure	30
2.2.4	Electrophysiological Recordings	31
2.3	Data Analysis	32
2.3.1	Behavioral Data	32
2.3.2	ERP Data	32
2.3.3	Classification	32
2.4	Statistical Analysis	33
2.4.1	Behavioral Data	33
2.4.2	ERP Data	33
2.4.3	Classification	34
2.5	Results	35
2.5.1	Behavioral Data	35
2.5.2	ERP Data	36
2.5.3	Classification	37
2.6	Discussion	38



2.7	Length Influence Experiment. . . . .	39
2.7.1	Introduction . . . . .	40
2.7.2	Methods . . . . .	40
2.7.3	Statistical Analysis. . . . .	41
2.7.4	Results. . . . .	41
2.7.5	Discussion . . . . .	43
2.8	Chapter Summary . . . . .	43
<b>3</b>	<b>ERPs and Quality Ratings Evoked by Word Stimuli and Varying Bit Rate Conditions . . . . .</b>	<b>45</b>
3.1	Introduction . . . . .	45
3.2	Methods . . . . .	46
3.2.1	Participants . . . . .	46
3.2.2	Material . . . . .	46
3.2.3	Experimental Design and Procedure. . . . .	47
3.2.4	Electrophysiological Recordings . . . . .	47
3.2.5	Data Analysis . . . . .	48
3.3	Statistical Analysis. . . . .	48
3.4	Results . . . . .	48
3.4.1	Behavioral Data. . . . .	48
3.4.2	P300 . . . . .	49
3.4.3	Classification . . . . .	49
3.5	Discussion . . . . .	51
3.6	Chapter Summary . . . . .	52
<b>4</b>	<b>ERPs and Quality Ratings Evoked by Sentence Stimuli at Different Reverberation Levels . . . . .</b>	<b>53</b>
4.1	Introduction . . . . .	53
4.2	Methods . . . . .	53
4.2.1	Participants . . . . .	53
4.2.2	Material . . . . .	54
4.2.3	Experimental Design . . . . .	54
4.2.4	Electrophysiological Recordings . . . . .	55
4.3	Results . . . . .	56
4.3.1	Behavioral and Subjective Data. . . . .	56
4.3.2	P300 . . . . .	59
4.4	Discussion . . . . .	61
4.5	Chapter Summary . . . . .	62
<b>5</b>	<b>EEG Frequency Band Power Changes Evoked by Listening to Audiobooks at Different Quality Levels . . . . .</b>	<b>63</b>
5.1	Introduction . . . . .	63
5.2	Methods . . . . .	64

- 5.2.1 Participants . . . . . 64
- 5.2.2 Materials . . . . . 65
- 5.2.3 Experimental Design . . . . . 66
- 5.2.4 Electrophysiological Recordings . . . . . 66
- 5.2.5 Statistical Analysis. . . . . 67
- 5.3 Results . . . . . 68
  - 5.3.1 Quality Ratings . . . . . 68
  - 5.3.2 Frequency Band Power Analysis . . . . . 68
- 5.4 Discussion . . . . . 71
- 5.5 Chapter Summary . . . . . 72
  
- 6 EEG Frequency Band Power Changes Evoked by Listening to Audiobooks with Varying Quality Profiles . . . . . 73**
  - 6.1 Introduction . . . . . 73
  - 6.2 Methods . . . . . 74
    - 6.2.1 Participants . . . . . 74
    - 6.2.2 Stimuli . . . . . 74
    - 6.2.3 Experimental Design . . . . . 75
    - 6.2.4 Electrophysiological Recordings . . . . . 76
    - 6.2.5 Statistical Analysis. . . . . 76
  - 6.3 Results . . . . . 78
    - 6.3.1 Instrumental Quality Estimation. . . . . 78
    - 6.3.2 Frequency Band Power Analysis . . . . . 78
  - 6.4 Discussion . . . . . 79
  - 6.5 Chapter Summary . . . . . 80
  
- 7 General Discussion and Future Work . . . . . 81**
  - 7.1 General Discussion . . . . . 81
    - 7.1.1 Phoneme Experiment . . . . . 82
    - 7.1.2 Word Experiment . . . . . 83
    - 7.1.3 Sentence Experiment . . . . . 83
    - 7.1.4 Audiobooks Experiment: Constant Quality . . . . . 84
    - 7.1.5 Audiobooks Experiment: Quality Profiles . . . . . 84
    - 7.1.6 Considerations on the Quality Formation Process. . . . . 84
  - 7.2 Practical Guidelines . . . . . 86
  - 7.3 Future Work . . . . . 87
  
- References . . . . . 91**

# Acronyms

2AFC	Two-alternative forced choice
ACR	Absolute category rating
ANOVA	Analysis of variance
AUC	Area under the curve
BCI	Brain-computer interface
CCR	Comparison category rating
CQS	Continuous quality scale
DCR	Degradation category rating
DSCQS	Double-stimulus continuous quality scale
EEG	Electroencephalography
ERP	Event-related potential
EU	European Union
FBP	Frequency band power
fMRI	Functional magnetic resonance imaging
ICT	Information and communication technology
ISI	Interstimulus interval
ITU	International Telecommunication Union
ITU-T	International Telecommunication Union—Telecommunication Standardization Sector
LDA	Linear discriminant analysis
MEG	Magnetoencephalography
MMN	Mismatch negativity
MOS	Mean opinion score
MUSHRA	Multi stimulus test with hidden reference and anchor
NS	Not significant
OP	Opinion test
POLQA	Perceptual objective listening quality assessment
QoE	Quality of experience

ROC	Receiver operating characteristic
SAM	Self-assessment manikin
SNR	Signal-to-noise ratio
SVM	Support vector machines