

Preface

The automatic detection of people's identity from their voices is part of modern telecommunication services. This generally requires the telephone transmission of speech to remote servers that perform the recognition task. The transmission may introduce severe distortions that degrade the system performance and hence represents one of the major challenges speech technologies are currently facing. Similarly, humans also cope with the difficulty of reliably identifying talkers from speech transmitted over communication channels, particularly if the utterance heard is of short duration.

This book addresses the evaluation of the human and of the automatic speaker recognition performances under different channel distortions caused by bandwidth limitation, codecs, and electroacoustic user interfaces, among other impairments. The main contribution of this work is the demonstration of the benefits of communication channels of extended bandwidth, together with an insight into how speaker-specific characteristics of speech are preserved through different transmissions. This book intends to assist students, researchers, and engineers to assess the speaker recognition performance employing transmitted speech. Particularly interesting for network engineers, this work provides sufficient motivation for considering speaker recognition as a criterion for the migration from narrowband to enhanced bandwidths, such as wideband and super-wideband.

This book was written in the context of my Ph.D. project, a research agreement between the Telekom Innovation Laboratories and the Technische Universität Berlin (TU Berlin), Germany, and the University of Canberra (UC), Australia. It therefore involved periods of research in Berlin, at the Quality and Usability Lab of the TU Berlin, and in Canberra, at the Human-Centred Computing Laboratory of the UC.

This work would not have been possible without the contributions of many supporters. I am most thankful to my two main supervisors Prof. Michael Wagner and Prof. Sebastian Möller for their support, direction, and advice throughout the course of my Ph.D. Their insightful comments and feedback have repeatedly guided my research focus towards its final form. My gratitude goes also to my third

supervisor Associate Professor Roland Göcke, for his constructive comments and input on my work. I wish to express my appreciation to the University of Canberra for the research grant award and to Deutsche Telekom AG for full project funding.

I would like to thank, among my colleagues in Berlin, the groups of speakers and listeners who volunteered to be recorded and to participate in my auditory tests. This was a kind favour as I required test participants who knew each another and had been exposed for a long period to one another's voices. Special thanks to Marcel Wältermann and to Janto Skowronek for their assistance with the audio transmission and reception scenarios of the auditory tests I conducted. I would also like to acknowledge Friedemann Köster's help in estimating quality of speech signals employing instrumental speech quality measures.

After having acquired a background on human voice perception and speaker recognisability, I was still new to automatic speaker recognition procedures. I would like to thank my colleague David Vandyke for the quick introduction to speaker recognition methodology when I arrived to Australia, jointly with Prof. Michael Wagner, which saved me considerable time.

Last, but not least, thanks to my family for their support, understanding, and patience while I was 17,800 km away from home (the current 2,300 km seems to be a bearable distance), and thanks to my friends, for their always-comforting words. Among other incentives, this encouraged my continuous work and dedication towards the completion of this book.

Berlin, June 2015

Laura Fernández Gallardo



<http://www.springer.com/978-981-287-726-0>

Human and Automatic Speaker Recognition over
Telecommunication Channels

Fernández Gallardo, L.

2016, XII, 169 p., Hardcover

ISBN: 978-981-287-726-0