MODELLING AUDITORY PROCESSING AND ORGANISATION

Distinguished Dissertations in Computer Science

Edited by
C.J. van Rijsbergen, University of Glasgow

The Conference of Professors of Computer Science (CPCS) in conjunction with the British Computer Society (BCS), selects annually for publication up to four of the best British Ph.D. dissertations in computer science. The scheme began in 1990. Its aim is to make more visible the significant contribution made by Britain - in particular by students - to computer science, and to provide a model for future students. Dissertations are selected on behalf of CPCS by a panel whose members are:

M. Clint, Queen's University, Belfast
R.J.M. Hughes, University of Glasgow
R. Milner, University of Edinburgh (Chairman)
K. Moody, University of Cambridge
M.S. Paterson, University of Warwick
S. Shrivastava, University of Newcastle upon Tyne
A. Sloman, University of Birmingham
F. Sumner, University of Manchester

MODELLING AUDITORY PROCESSING AND ORGANISATION

Martin Cooke
University of Sheffield

CAMBRIDGE
UNIVERSITY PRESS

PUBLISHED BY THE PRESS SYNDICATE OF THE UNIVERSITY OF CAMBRIDGE
The Pitt Building, Trumpington Street, Cambridge, United Kingdom

CAMBRIDGE UNIVERSITY PRESS
The Edinburgh Building, Cambridge CB2 2RU, UK
40 West 20th Street, New York NY 10011–4211, USA
477 Williamstown Road, Port Melbourne, VIC 3207, Australia
Ruiz de Alarcón 13, 28014 Madrid, Spain
Dock House, The Waterfront, Cape Town 8001, South Africa

http://www.cambridge.org

First published 1993
First paperback edition 2005

A catalogue record for this book is available from the British Library

ISBN 0 521 45094 2 hardback
ISBN 0 521 61938 6 paperback

for Jo, Sam and Anna

Contents

Acknowledgments

It is a pleasure to thank all those individuals whose research has been influential in shaping this work. The strong auditory community of the UK is sufficiently cohesive to provide great help to anyone setting out on a modelling study such as the one described here. I have enjoyed productive discussions with research groups at the Universities of Cambridge (Roy Patterson and John Holdsworth), Edinburgh (Mark Terry), Keele (Bill Ainsworth and Ted Evans), Kiel (Michel Scheffers), Loughborough (Ray Meddis and Michael Hewitt), Nottingham (Quentin Summerfield and Richard Stubbs), Southampton (Bob Damper) and Sussex (Chris Darwin, Valter Ciocca and Roy Gardner), and with Roger Moore of the Speech Research Unit, RSRE.

I found collaboration with members of the Laboratoire D'Informatique pour la Mecanique et les Sciences de l'Ingenieur in Paris (Christophe D'Alessandro, Jean-Sylvain Lienard and Maxine Eskenazi) in the 1989 Cochleagram-Reading Workshop stimulating; in particular, the use of resynthesis was prompted by the work of Jean-Sylvain Lienard.

My interest in auditory modelling was inspired by David Schofield, a colleague of earlier days at the National Physical Laboratory, whose work in matching physiological and psychophysical accounts of frequency selectivity has, I believe, proved catalytic for later developments of the gammatone filterbank.

Thanks to: Steve Beet and Robin Sharpe for providing a good deal more insight into filter design and signal processing than I could have obtained from textbooks alone; Chris Darwin for the /ru/-/li/ stimulus used to illustrate grouping in Chapter 5; to the Centre for Speech Technology Research at Edinburgh University and the Department of Linguistics and Phonetics at Leeds for some of the utterances used; and to Lori Lamel for the loan of TIMIT.

Locally, I have enjoyed working with members of SPLASH, the SPeech LAboratory at SHeffield; particular credit must go to Malcolm 'disc-crasher' Crawford, who, apart from many stimulating discussions and all things confectionery, forced me to hit the return key on many crucial processes, when I preferred to pause contemplatively for a few minutes before discovering that an idea was not going to work. Thanks also to Guy Brown for proof-reading beyond the call and useful suggestions. Other members of the Department of Computer Science have, with a mixture of ribaldry and good-humour, made Sheffield a great place to work.

Special thanks are due to my supervisor, Phil Green, who had this work in mind from as long ago as 1984. His encouragement and confidence helped tremendously throughout.

The material presented in chapter 3 is re-expressed from a paper published in the journal Computer Speech and Language and is used with permission from Academic Press Ltd.