Interactive Auditory Demonstrations

Martin Cooke
Speech and Hearing Research
Department of Computer Science
University of Sheffield
Regent Court
211 Portobello Street
Sheffield S1 4DP
UK

m.cooke@dcs.shef.ac.uk

Abstract

The subject matter of speech and hearing is packed full of phenomena and processes which lend themselves to or require auditory demonstration. In the past, this has been achieved through passive media such as tape or CD (e.g. Houtsma et al, 1987; Bregman & Ahad, 1995). The advent of languages such as MATLAB which supports sound handling, modern interface elements and powerful signal processing routines, coupled with the availability of fast processors and ubiquitous soundcards allows for a more interactive style of demonstration. A significant effort is now underway in the speech and hearing community to exploit these favourable conditions (see the MATISSE proceedings (1999), for instance).

Excitingly, it is now possible to allow exploratory access to part or all of the parameter space underlying each phenomenon. Over the past 18 months, more than 20 interactive auditory demonstrations have been produced at Sheffield as part of an ongoing project to provide teaching material for the diverse disciplines which contribute to speech and hearing. Many of the demonstrations are suitable for undergraduate courses, while others encode phenomena which are primarily of interest to researchers.

The motivation for and design ethos behind this project has been described previously in Cooke & Brown (1999) and Wrigley, Cooke & Brown (1999). In this extended abstract, a gallery of screenshots which focus on the auditory (as opposed to speech) demonstrations is provided. The aim is to show the breadth of what is possible in a relatively short time and to encourage others to produce similar tools.

The demonstrations can be freely downloaded via http://www.dcs.shef.ac.uk/~martin.

Acknowledgements

Demonstrations described here were programmed by Guy Brown, Martin Cooke and Stuart Wrigley (Sheffield, UK) and Dan Ellis (ICSI, Berkeley, USA). Stuart Cunningham and Ljubomir Josifovski helped with the testing. Funding for some of the development work was provided by the ELSNET LE Training Showcase, 98/02.

References

Assmann, P.F. and Summerfield, Q. (1990) Modelling the perception of concurrent vowels: vowels with different fundamental frequencies. *Journal of the Acoustical Society of America*, 88(2), 680-697.

Barker, J. and Cooke, M.P. (1999) Is the sine-wave speech cocktail party worth attending? *Speech Communication*, 27(3-4), 159-174.

Bloothooft, G., van Dommelen, W. , Espain, C., Hazan, V. , Huckvale, M. and Wigforss, E., *The landscape of future education in speech communication sciences* (Utrecht Institute of Linguistics Publications ISBN 90-5434-069-X, 1998).

Bregman, A.S. and Ahad, P., *Demonstrations of auditory scene analysis: the perceptual organisation of sound*, CDROM, MIT Press, Cambridge, Mass, 1995.

Cooke, M.P. & Brown, G.J. (1999) Interactive explorations in speech and hearing, *Journal of the Acoustical Society of Japan*, in press.

Cooke, M.P. (1991). *Modelling auditory processing and organisation.* PhD Thesis. Published by Cambridge University Press, 1993.

Hartmann, W.M. & Johnson, D. (1991). Stream segregation and peripheral channeling, *Music Perception*, 9(2), 155-184.

Houtsma, A.J.M., Rossing, T.D. and Wagenaars, W.M. (1987) *Auditory Demonstrations Compact Disc.* Available through the Acoustical Society of America.

Moore, B.C.J., Glasberg, B.R. & Peters, R.W. (1985), Relative dominance of individual partials in determining the pitch of complex tones, *Journal of the Acoustical Society of America*, 77, 1853-1860.

Moore, B.C.J. (1997) *An Introduction to the Psychology of Hearing (4th ed)*, Academic Press.

van Noorden, L.P.A.S. (1975) Temporal coherence in the perception of tone sequences, Ph.D. Thesis, Eindhoven University of Technology.

MATISSE (1999) *Methods and Tool Innovations in Speech Science Education*, London, April.

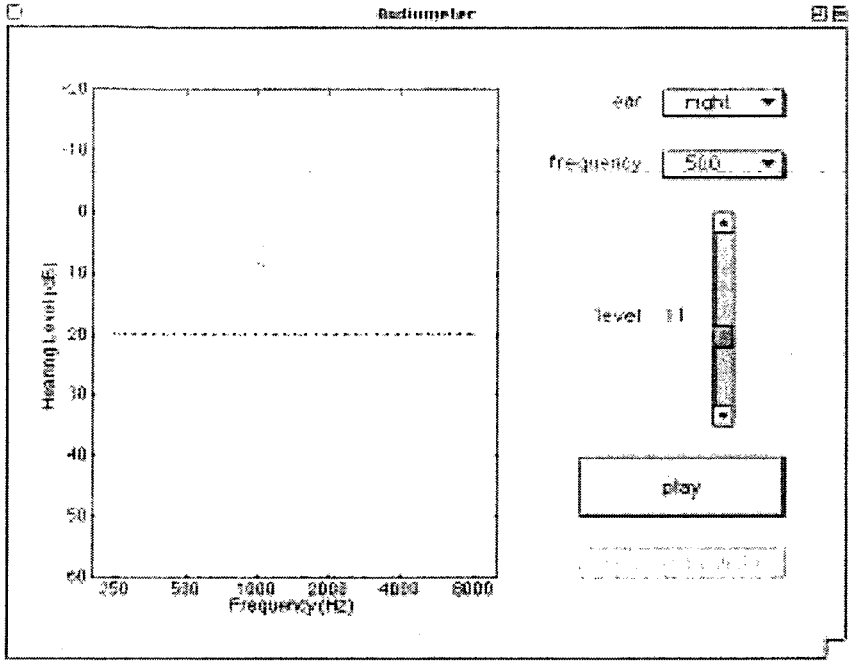
Remez, R.E., Rubin, P.E., Pisono, D.B. and Carrell, T.D. (1981). Speech perception without traditional speech cues, *Science*, 212, 947-950.

Patterson & Holdsworth (1990). In: *Adv. in Speech, Hearing & Language* Proc., Vol. 3 (ed: Ainsworth, W.), JAI Press.

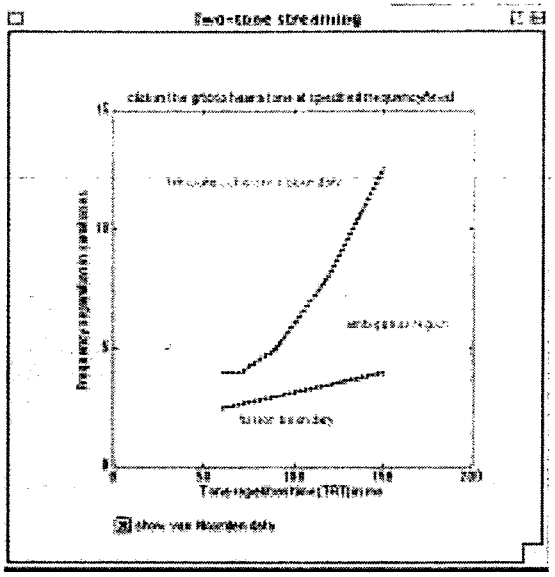
Scheffers, M.T.M. (1983) *Sifting vowels: auditory pitch analysis and sound segregation*, PhD Thesis, Groningen University.

Warren, R.M. (1970). Perceptual restoration of missing speech sounds *Science*, 167, 392-393.

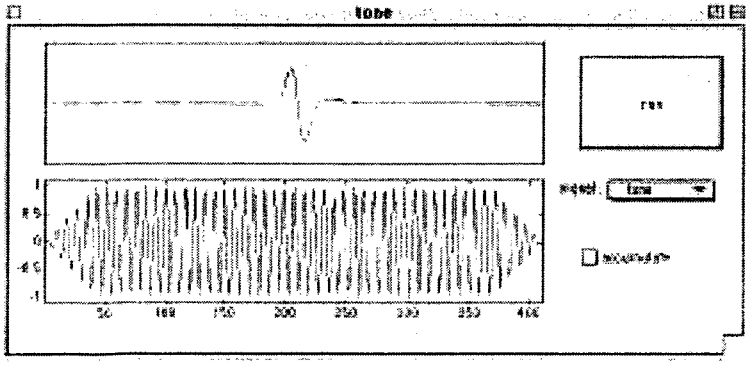
Wrigley, S., Cooke, M.P. & Brown, G.J. (1999) *Interactive learning in speech and hearing*, Proc. MATISSE Workshop.



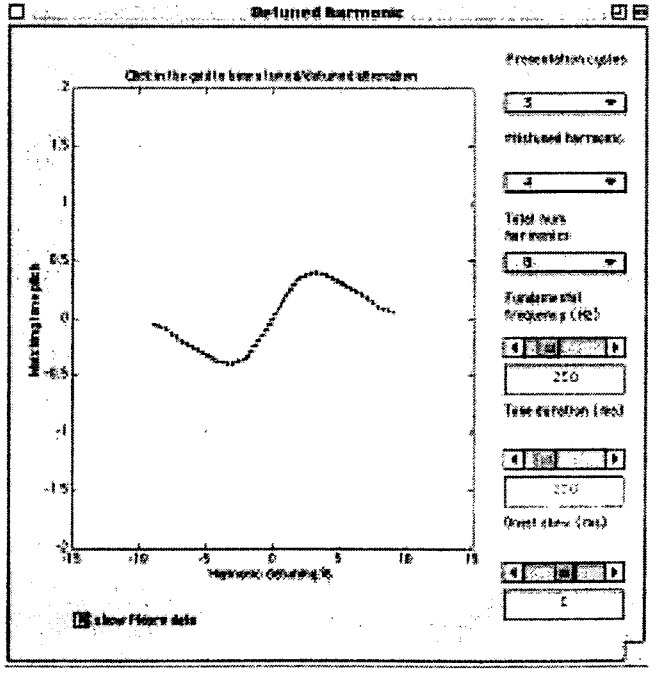
audiometer - a training tool for pure-tone audiometry



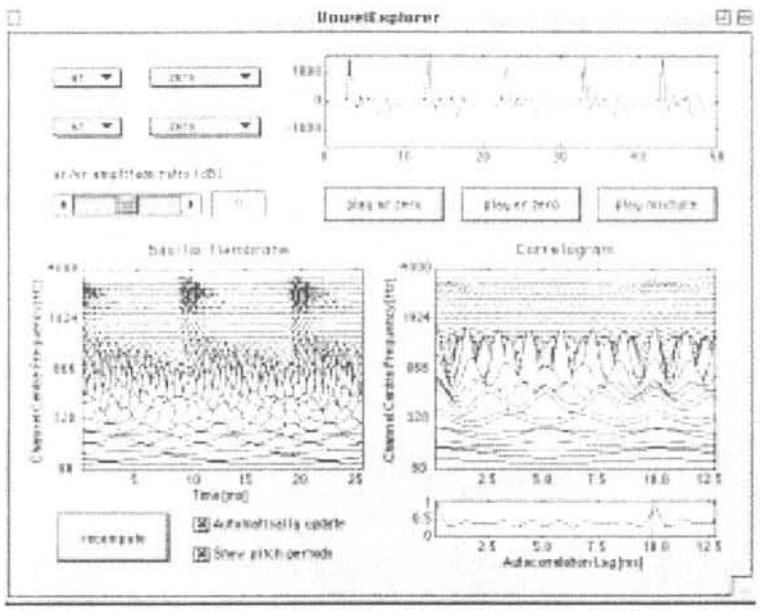
streamer - segregation and fusion of alternating tones (Van Noorden,



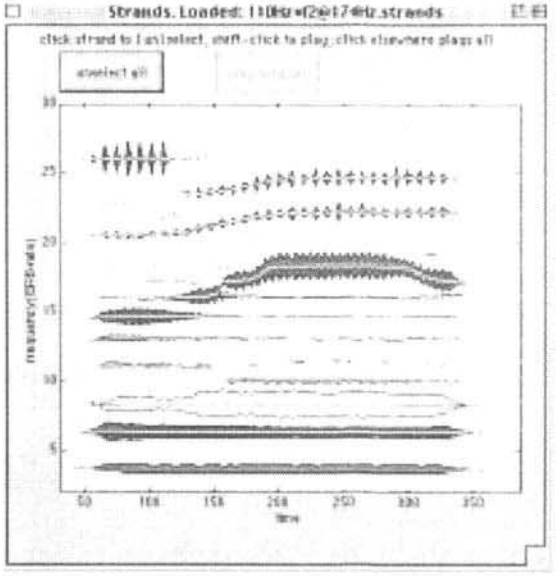
bm - animation of a passive, linear model of basilar membrane motion



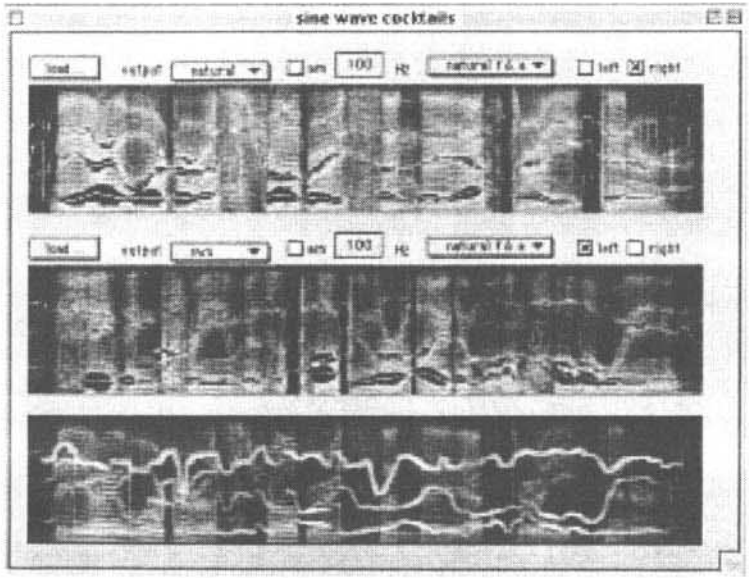
detuning - effect of a mistuned harmonic on perceived pitch and number of sources (Moore et al, 1985)



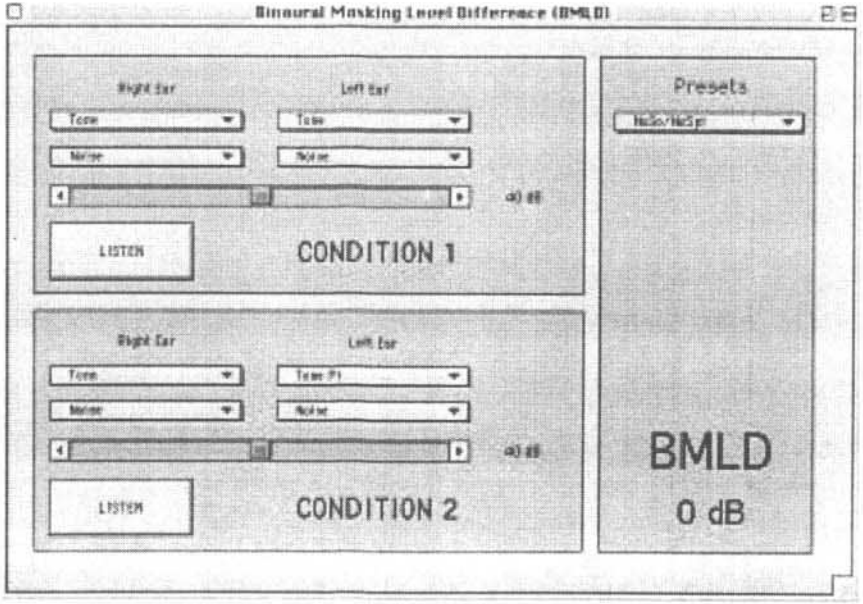
vowelExplorer - perception and identification of concurrent vowels (Scheffers, 1983)



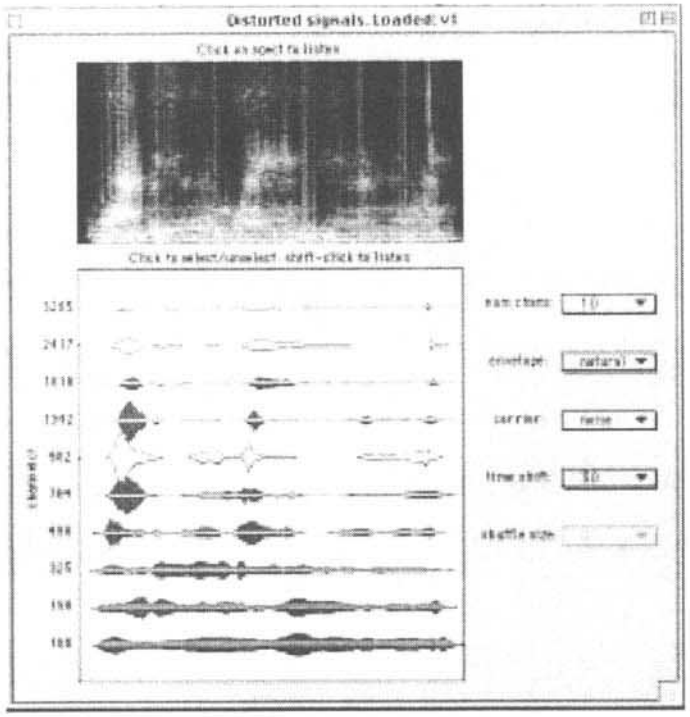
strands - auditorily-motivated time-frequency representation for analysis and resynthesis (Cooke, 1991)



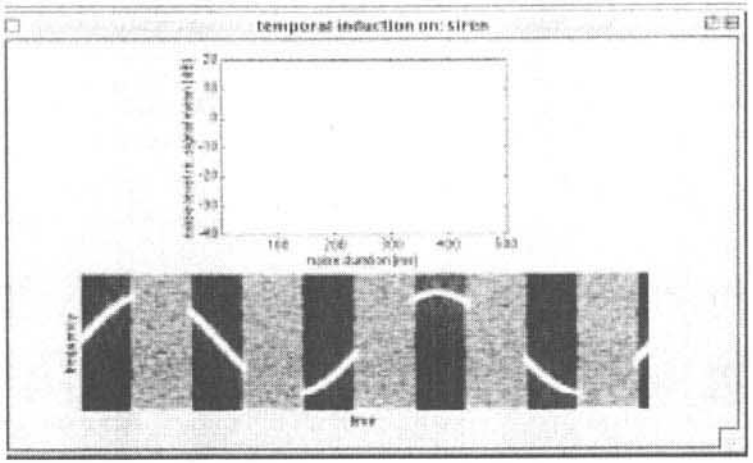
sws - cocktail party sine-wave speech (Barker & Cooke, 1999; Remez et al, 1981)



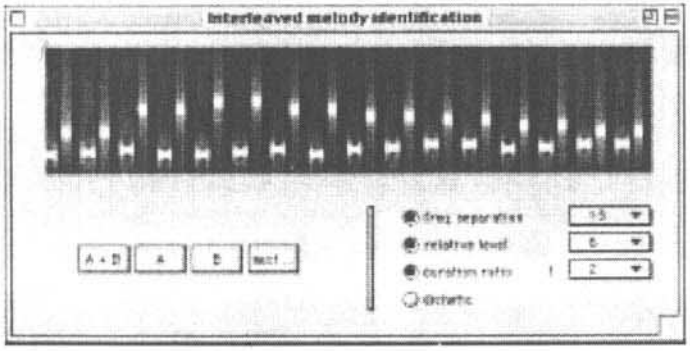
bmld - binaural masking level difference (see Moore, 1997)



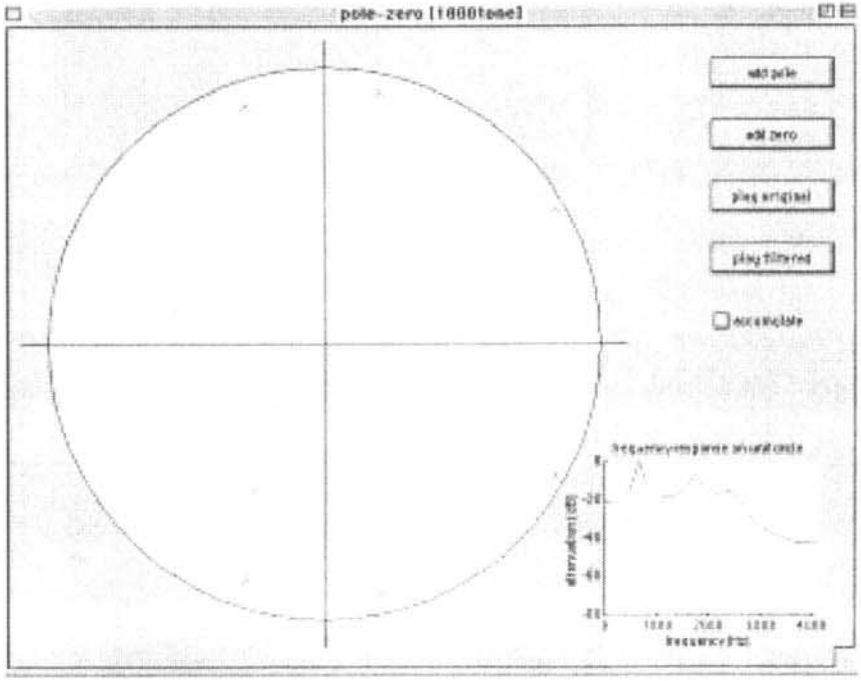
distortion - perceptual effect of spectral and temporal distortions



ti - temporal auditory induction (Warren, 1970)



intmel - segregation of interleaved melodies (Hartmann & Johnson, 1991)



polezero - effect of pole/zero placement