

Publications 2002



Dissertations

Engwall O. *Tongue Talking - Studies in Intraoral Speech Synthesis*. Department of Speech, Music and Hearing, KTH, Stockholm. TRITA-TMH 2002: 4, ISSN 1104-5787, ISBN 91-7283-258-9.*

Guettler, Knut. *The Bowed String. On the Development of Helmholtz Motion and On the Creation of Anomalous Low Frequencies*. Department of Speech, Music and Hearing, KTH. TRITA-TMH 2002:3, ISSN 1100-5787, 91-7283-279-7.

Gustafson J. *Developing multimodal spoken dialogue systems. Empirical studies of spoken human-computer interaction*. Department of Speech, Music and Hearing, KTH, Stockholm. TRITA-TMH 2002:8, ISSN 1104-5787.*

Megyesi B. *Data-Driven Syntactic Analysis - Methods and Applications for Swedish*. Department of Speech, Music and Hearing, KTH, Stockholm. TRITA-TMH 2002:7, ISSN 1104-5787, ISBN 91-7283-384-X.*

Licentiate Thesis

Magnuson T (2002). Assessment of writing difficulties and evaluation of computerized writing support. Department of Speech, Music and Hearing, KTH, Stockholm. TRITA-TMH 2002:6, ISSN 1104-5787, ISBN 91-7283-375-0.

Journal Publications

Friberg A. & Sundström A. (2002). Swing ratios and ensemble timing in jazz performance: Evidence for a common rhythmic pattern. (Invited paper). *Music Perception* 19/3: 333-349.

Juslin PN, Friberg A & Bresin R (2002). Toward a computational model of expression in performance: The GERM model. *Musicae Scientiae special issue 2001-2002*: 63-122.

Leijon A (2002). Estimation of auditory information transmission capacity using a hidden Markov model of speech stimuli. *Acustica* 88/3: 423-432.

Megyesi B (2002). Shallow parsing with pos taggers and linguistic knowledge. *JMLR, Journal of Machine Learning Research: Special Issue on Shallow Parsing*. (2): 639-668.*

Rocchesso D, Fernström M, Moynihan B & Bresin R (2002). The sounding object. Flash presentation included in the premier CG&A cd-rom. *IEEE Computer Graphics*, 22:4.

Schoonderwaldt E, Friberg A, Bresin R & Juslin PN (2002). A system for improving the communication of emotion in music performance by feedback learning. *Journal of the Acoustical Society of America* 111/5: 2471.

Södersten M, Granqvist S, Hammarberg B & Szabo A (2002). Vocal behaviour and vocal loading factors for preschool teachers at work studied with binaural DAT recordings. *Journal of Voice* 16/3: 356-371.

Ternström S, Södersten M, Bohman M (2002). Cancellation of simulated environmental noise as a tool for measuring vocal performance during noise exposure. *Journal of Voice* 16/2: 195-206.

Books, Book chapters

Beskow J, Granström B & House D (2002). A multimodal speech synthesis tool applied to audio-visual prosody. In: Keller E, Bailly G, Monaghan A, Terken J & Huckvale M, eds. *Improvements in Speech Synthesis*. New York: John Wiley & Sons, Inc, 372-382.*

Bresin R & Hansen K (2002). Multimodal perception of model-based rolling. In: *Sounding Objects Deliverable No 10*, Perceptual Assessment of Models.

Friberg A & Battel GU (2002). Structural communication. In: Parncutt R & McPherson GE (Eds.) *The Science and Psychology of Music Performance: Creative Strategies for Teaching and Learning*. New York: Oxford University Press; 199-218.

Gustafson K & House D (2002). Prosodic parameters for a "fun" speaking style. In: Keller E, Bailly G, Monaghan A, Terken J & Huckvale M (Eds). *Improvements in Speech Synthesis*. New York: John Wiley & Sons, 264-272.*

Granström B, House D & Beskow J (2002). Speech and gestures for talking faces in conversational dialogue systems. In: Granström B, House D & Karlsson I (eds). *Multimodality in language and speech systems*. Dordrecht: Kluwer Academic Publishers, 209-241.*

Granström B, House D & Karlsson I (eds.) (2002). *Multimodality in Language and Speech Systems*. Dordrecht: Kluwer Academic Publishers.*

Hansen KF (2002). Music besides grooves. In: Depraetere F & Willemsen P (eds): *Pitch - Mutating Turntables: Argos Festival Catalogue 2002*. (Argos Editions, ISBN 90-76855-13-7), 136-146.

Hansen K & Bresin R (2002). Scratching: From analysis to modelling. In: *Sounding Objects Deliverable No 8*, Models and Algorithms for Control of Sounding Objects.

House D (2002). The interaction of pitch range and temporal alignment in the perception of interrogative mode in Swedish (Abstract). In: Hawkins S & Nguyen N (Eds.). *Temporal Integration in the Perception of Speech*, University of Cambridge.*

House D & Granström B (2002). Multimodal speech synthesis: Improving information flow in dialogue systems using 3D talking heads. In: Scott D (ed.). *Artificial Intelligence: Methodology, Systems, and Applications*, 10th International Conference, AIMS 2002, Springer-Verlag, Berlin. 1-10.*

Jansson E (2002). *Acoustics for Violin and Guitar Makers*. Compendium, 4th edition. Dept of Speech, Music and Hearing, KTH, Sweden.

Ternström S & Karna DR (2002). Choir. In: Parncutt R, McPherson G (Eds). *The Science and Psychology of Music Performance*. Chapter 17. New York: Oxford University Press, 2002.

Conference Proceedings

Andersen SV, Kleijn BW, Hagen R, Linden J, Murthi MN & Skoglund J (2002). iLBC - A linear predictive coder with robustness to packet losses. In: *Proc IEEE Workshop on Speech Coding*, 2002, Tsukuba.

Backstrom T, Alku P & Kleijn WB (2002). A time-domain reformulation of linear prediction equivalent to the LSP decomposition. In: *Proc IEEE Int Conf Acoust Speech Sign Process*, I: 661-664.

Beskow J, Edlund J & Nordstrand M (2002). Specification and realisation of multimodal output in dialogue systems. *Proc of ICSLP 2002. Denver, Colorado, USA, Sept 17-20*; I: 181-184.*

Beskow J, Granström B & Spens K-E (2002). Articulation strength - Readability experiments with a synthetic talking face. *Proc of Fonetik 2002, May 29-31, Stockholm. TMH-QPSR, KTH*, Vol 44: 97-100.

Bohman M, Ternström S & Södersten M (2002). The use of channel estimation techniques for investigating vocal stress in noisy work environments. In: *Proc of Baltic-Nordic Acoustical Meeting*, Copenhagen, Denmark, August 2002.

Bresin R & Nettheim N (2002). Displaying the expressive parameters of music performance. In: *Proc of ICMPC7 - 7th International Conference on Music Perception & Cognition*, Sydney, July 17-21; p. 6.

- Bresin R (2002). Color feedback in expressive music performance. In: *Proc of ICMPC7 - 7th International Conference on Music Perception & Cognition*, Sydney, July 17-21; p 14.
- Bresin R (2002). Importance of note-level control in automatic music performance. In: *Proc of RENCON 2002 - performance RENDERing piano CONtest*, Kyoto, 1-6.
- Bresin R, Friberg A & Sundberg J (2002). Director musices: The KTH performance rules system. In: *Proc of SIGMUS-46*, Kyoto, 43-48.
- Carlson R, Granström B, Heldner M, House D, Megyesi B, Strangert E & Swerts M (2002). Boundaries and groupings - the structuring of speech in different communicative situations: a description of the GROG project. *Proc of Fonetik 2002, May 29-31, Stockholm. TMH-QPSR, KTH*, Vol 44: 65-68.
- Carlson R, Sjölander A & Sigvardson T (2002). Data-driven formant synthesis. *Proc of Fonetik 2002, May 29-31, Stockholm. TMH-QPSR, KTH*, Vol 44: 121-124.
- Cerrato L (2002). A comparison between feedback strategies in Human-to-Human and Human-Machine communication. In: *Proc of ICSLP-2002, Denver, Colorado, USA, Sept 17-20*; I: 557-560.
- Cerrato L & Ekeklint S (2002). Different ways of ending human-machine interactions. *Proc of AAMAS02 Workshop*, Bologna.*
- Cerrato L & Skhiri M (2002). Quantifying non-verbal communicative behaviour in face-to-face human dialogues. Abstract with poster presented at *First Pan-American/Iberian Meeting on Acoustics Cancun*, Mexico, Dec 2-6, 2002.
- Chollet G, Faundez-Zanuy M, Kleijn B, Kubin G, McLaughlin S & Petek B (2002). Future work in COST action 277 on nonlinear speech processing. In: *Proc of EUSIPCO 2002*, Toulouse.
- Dahl S & Granqvist S (2002). A new method for estimating internal drift and just noticeable difference in perception of continuous tempo drift. Poster presented at *The Neurosciences and Music, Mutual interactions and implications on developmental functions*. Fondazione Pierfranco e Luisa Mariani, Venice, San Servolo Island, Oct 25-27, 2002.
- Edlund J & Nordstrand M (2002). Turn-taking gestures and hour-glasses in a multi-modal dialogue system. *Proc of ISCA Workshop Multi-Modal Dialogue in Mobile Environments*. Kloster Irsee, Germany, June 17-19, 2002.*
- Edlund J, Beskow J & Nordstrand M (2002). GESOM - A model for describing and generating multi-modal output. *Proc of ISCA Workshop on Multi-Modal Dialogue in Mobile Environments*, Kloster Irsee, Germany, June 17-19, 2002.
- Elenius D & Blomberg M (2002). Characteristics of a low reject mode speaker verification system. In: *Proc of ICSLP 2002*, Denver, Colorado, USA, Sept 17-20; 1385-1388.*
- Engwall O (2002). Evaluation of a system for concatenative articulatory visual speech synthesis. In: *Proc of ICSLP 2002* Denver, Colorado, USA, Sept 17-20; 665-668.*
- Fant G, Kruckenberg A, Gustafson K & Liljencrants J (2002). A new approach to intonation analysis and synthesis of Swedish. *Proc of Fonetik 2002, May 29-31, Stockholm. TMH-QPSR, KTH*, Vol 44: 161-64.
- Fant G, Kruckenberg A, Gustafson K, Liljencrants J & Botinis A (2002). Individual variations in prominence correlates. Some observations from "lab-speech". *Proc of Fonetik 2002, May 29-31, Stockholm. TMH-QPSR, KTH*, Vol 44: 177-180.
- Faundez-Zanuy M, Nilsson M & Kleijn WB (2002). On the relevance of bandwidth extension for speaker verification. In: *Proc of ICSLP'2002, Int Conf Spoken Lang Process 2002, Denver, Colorado, USA, Sept 17-20*; 2317-2320.
- Faundez-Zanuy M, Nilsson M & Kleijn WB (2002). On the relevance of bandwidth extension for speaker identification. In: *Proc EUSIPCO, 2002*, Toulouse.
- Friberg A, Schoonderwaldt E, Juslin PN & Bresin R (2002). Automatic extraction of expressive cues for cognitive feedback. In: *Proc of ICMPC7 - 7th International Conference on Music Perception & Cognition*, Sydney, July 17-21; p. 14.
- Friberg A, Schoonderwaldt E, Juslin PN & Bresin R (2002). Automatic real-time extraction of musical expression. In: *Proc of the*

International Computer Music Conference - ICMC 2002, Göteborg; 365-367.

Granström B, House D & Swerts MG (2002). Multimodal feedback cues in human-machine interactions. In: Bel B & Marlien I (eds.). *Proc of the Speech Prosody 2002 Conference*, April 11-13, 2002. Aix-en-Provence: Laboratoire Parole et Langage; 347-350.*

Gustafson J & Sjölander K (2002). Voice transformations for improving children's speech recognition in a publicly available dialogue system. In: *Proc of ICSLP 2002, Denver, Colorado, USA, Sept 17-20*; I: 297-300.*

Gustafson J, Bell L, Boye J, Edlund J & Wiren M (2002). Constraint manipulation and visualization in a multimodal dialogue system. In: *Proc of the ISCA Workshop Multi-Modal Dialogue in Mobile Environments*. Kloster, Irsee, Germany, June 17-21.*

Gustafson-Capkova S & Megyesi B (2002). Silence and discourse context in read speech and dialogues in Swedish. In: Bernard Bel & Isabelle Marlien (Eds). *Proc of Speech Prosody 2002 Conference*, Aix-en-Provence: Laboratoire Parole et Langage. ISBN 2-9518233-0-4. pp 363-366.*

Hincks R (2002). Speech recognition for language teaching and evaluating: a study of existing products. *Proc of ICSLP 2002, Denver, Colorado, USA, Sept 17-20*, 2: 733-736.

House D (2002). Intonational and visual cues in the perception of interrogative mode in Swedish. In: *Proc of ICSLP 2002, Denver, Colorado, USA, Sept 17-20*; 3: 1957-1960.*

Hunnicut S & Magnuson T (2002). Sentences for symbol users: email and echat. In: *Proc of ISAAC '02, 10th Biennial Conference of the International Society for Augmentative and Alternative Communication*, Odense, Denmark; 467-468.

Jansson EV & Molin N-E (2002). Fundamental violin acoustics related to quality. In: *Proc of the 32nd Int Acoustical Conf – EAA Symposium “Acoustics Banská Stiavnica 2002”*, Slovakia, Sept 10-12.

Jefremov A & Kleijn WB (2002). Spline-based continuous-time pitch estimation. In: *Proc IEEE Int Conf Acoust Speech Sign Process*, Orlando. I: 337-340.

Johansson M, Blomberg M, Elenius K, Hoffsten L-E & Torberger A (2002). Phoneme recognition for the hearing impaired. *Proc of Fonetik 2002, May 29-31, Stockholm. TMH-QPSR, KTH*, Vol 44: 109-112.*

Johansson M, Blomberg M, Elenius K, Hoffsten L-E & Torberger A (2002). A phoneme recognizer for the hearing impaired. In: *Proc of ICSLP'2002, Denver, Colorado, USA, Sept 17-20*; 433-436.*

Juslin PN, Friberg A & Bresin R (2002). Computational modelling of different aspects of expressivity: The GERM model. In: *Proc of ICMPC7 - 7th International Conference on Music Perception & Cognition*, Sydney, July 17-21; p. 13.

Karneböck S (2002): Expanded examinations of a low frequency modulation feature for speech/music discrimination. In: *Proc of ICSLP 2002, Denver, Colorado, USA, Sept 17-20*; 3: 2009-2012.*

Kim MY & Kleijn WB (2002). KLT-based classified VQ for the speech signal. In: *Proc IEEE Int Conf Acoustics, Speech, Signal Processing*, 2002, Orlando, 1: 645-648.

Kim MY & Kleijn WB (2002). KLT-based classified VQ with companding. In: *Proc. IEEE Workshop on Speech Coding, 2002*, Tsukuba, Japan. Pp 8-10.

Kleijn WB (2002). Enhancement of coded speech by constrained optimization. In: *Proc IEEE Workshop on Speech Coding*, Oct 6-9, 2002, Tsukuba, Japan.

Kleijn WB & Talkin D (2002). Compact speech representations for speech synthesis. In: *Proc IEEE Workshop on Speech Synthesis*, Sept 11-13, 2002, Santa Monica.

Laukka P, Bresin R & Lindström E (2002). The teacher's perspective on expression in music performance: Preliminary evidence from music conservatories. In: *Proc of ICMPC7 - 7th International Conference on Music Perception & Cognition*, Sydney, July 17-21; p. 13.

Lindström E, Juslin PN, Bresin R & Williamon A (2002). Students' views on expressivity in music performance: A questionnaire study. In: *Proc of ICMPC7 - 7th International Conference on Music Perception & Cognition*, Sydney, July 17-21; pp. 12-13.

Megyesi B & Carlson R (2002). Data-driven methods for building a Swedish Treebank. Presented at the Swedish Treebank Symposium, Nov 28-29, 2002, Växjö University, Sweden.*

Megyesi B & Gustafson-Capkova S (2002). Production and perception of pauses and their linguistic context in read and spontaneous speech in Swedish. In: *Proc of ICSLP'2002, Denver, Colorado, USA, Sept 17-20*; 3: 2153-2156.*

Nilsson M, Gustafsson H, Andersen SV & Kleijn WB (2002). Gaussian Mixture Model based Mutual Information estimation between Frequency Bands in Speech. In: *Proc IEEE Int Conf Acoust Speech Sign Process, 2002, Orlando, 1*: 525-528.

Nordqvist P & Leijon A (2002). Characterising the dynamic behaviour of advanced hearing instruments in real-life environments. Poster presented at the *International Hearing Aid Research (IHCON) Conference 2002, Lake Tahoe, CA*.

Pakucs B (2002). VoiceXML-based dynamic plug and play dialogue management for mobile environments. In: *Proc of the ISCA Workshop Multi-Modal Dialogue in Mobile Environments, Kloster, Irsee, Germany*.*

Rydin S (2002). Building a hyponymy lexicon with hierarchical structure. In: *Proc of the SIGLEX Workshop on Unsupervised Lexical Acquisition, ACL'02*, 26-33.*

Skantze G (2002). Coordination of referring expressions in multimodal human-computer dialogue. *Proc of ICSLP 2002, Denver, Colorado, USA, Sept 17-20*; 553-556.*

Smeds K, Grant F, Keidser G, Dillon H, Brew C & Zakis J (2002). Is normalisation of overall loudness an appropriate goal for hearing aid prescription? Preliminary results. In: *Proc of International Society of Audiology, Melbourne Australia, March 17-22, 2002*.

Smeds K, Keidser G, Dillon H, Grant F, Convery E & Zakis J (2002). Is normalisation of overall loudness an appropriate goal for hearing aid prescription? In: *Proc of International Hearing Aid Research Conference, IHCON 2002, Lake Tahoe, USA, August 21-25, 2002*.

Sundberg J (2002). Breathing and phonation. A review of some recent investigations (Invited

paper). *6th International Scientific Conference Research in Education and Rehabilitation Sciences, Zagreb, Oct 23-25, 2002. Published on Internet*.

Sundberg J (2002). Der menschliche Stimmausdruck - messbar oder nicht? (Invited paper). *4. Stuttgarter Stimmtagung, Oct 3-6, 2002. Published on Internet*.

Ternström S (2002). Applications of the analysis-by-synthesis paradigm in choir and voice acoustics research. *1st International Conference on the Physiology and Acoustics of Singing (PAS), Groningen, the Netherlands, Oct 2-5, 2002. (Published on CD-rom)*.

Vafin R & Kleijn WB (2002). Entropy-constrained polar quantization: Theory and an application to audio coding. In: *Proc IEEE Int Conf Acoust Speech Sign Process, 2002, Orlando, Florida, USA, 2*: 1837-1840.

Öster A-M, House D, Protopapas A & Hatzis (2002). A presentation of a new EU project for speech therapy: OLP (Ortho-Logo-Paedia). *Proc of Fonetik 2002, May 29-31, Stockholm. TMH-QPSR, KTH, Vol 44*: 45-48.*

TMH-QPSR, KTH Speech, Music and Hearing – Quarterly Progress and Status Report

From 2001 and on, TMH-QPSR is published on line only. Nevertheless, Proceedings from Fonetik 2002 was printed and published as Volume 44, 2002.

Birch P, Gümöes B, Prytz S, Karle A, Stavad H & Sundberg J (2002). Effects of a velopharyngeal opening on the sound transfer characteristics of the vowel [a]. *TFH-QPSR, KTH, Vol 43*: 9-15.
<http://www.speech.kth.se/qpsr/tmh/#V43>

Bretos J & Sundberg J (2002). Measurements of vibrato parameters in long sustained crescendo notes as sung by ten sopranos. *TFH-QPSR, KTH, Vol 43*: 37-44.
<http://www.speech.kth.se/qpsr/tmh/#V43>

Cerrato L (2002). Some characteristics of feedback expressions in Swedish. *Proc of Fonetik 2002, May 29-31, Stockholm. TMH-QPSR, KTH, Vol 44*: 101-104.

- Drioli C (2002). A flow waveform adaptive mechanical glottal model. *TFH-QPSR, KTH*, Vol 43: 69-80.
<http://www.speech.kth.se/qpsr/tmh/#V43>
- Granström B, Karlsson I & Spens K-E (2002). SYNFACE - a project presentation. *Proc of Fonetik 2002, May 29-31, Stockholm. TMH-QPSR, KTH*, Vol 44: 93-96.*
- Hincks R (2002). Speech synthesis for teaching lexical stress. *Proc of Fonetik 2002, May 29-31, Stockholm. TMH-QPSR, KTH*, Vol 44: 153-156.
- House D (2002). Perception of question intonation and facial gestures. *Proc of Fonetik 2002, May 29-31, Stockholm. TMH-QPSR, KTH*, Vol 44: 41-44.*
- Magnuson T & Hunnicutt S (2002). Measuring the effectiveness of word prediction: The advantage of long-term use. *TFH-QPSR, KTH*, Vol 43: 57-67.
<http://www.speech.kth.se/qpsr/tmh/#V43>
- Mürbe D, Pabst F, Hofmann G & Sundberg J (2002). Effects of a professional solo singer education on auditory and kinesthetic feedback - a longitudinal study of singers' pitch control. *TFH-QPSR, KTH*, Vol 43: 81-87.
<http://www.speech.kth.se/qpsr/tmh/#V43>
- Nordqvist P & Leijon A (2002). Automatic classification of the telephone listening environment in a hearing aid. *TFH-QPSR, KTH*, Vol 43: 45-49.
<http://www.speech.kth.se/qpsr/tmh/#V43>
- Sjölander K (2002). Recent developments regarding the WaveSurfer speech tool. *Proc of Fonetik 2002, May 29-31, Stockholm. TMH-QPSR, KTH*, Vol 44: 53-56.*
- Stone RE, Cleveland TF, Sundberg J & Prokop J (2002). Aerodynamic and acoustical measures of speech, operatic, and Broadway vocal styles in a professional female singer. *TFH-QPSR, KTH*, Vol 43: 17-29.
<http://www.speech.kth.se/qpsr/tmh/#V43>
- Sundberg J, Thalén M, Alku P & Vilkmán E (2002). Estimating perceived phonatory pressedness in singing from flow glottograms. *TFH-QPSR, KTH*, Vol 43: 89-96.
<http://www.speech.kth.se/qpsr/tmh/#V43>
- Ternström S (2002). Choir acoustics – an overview of scientific research published to date. *TFH-QPSR, KTH*, Vol 43: 1-8.
<http://www.speech.kth.se/qpsr/tmh/#V43>
- Zangger-Borch D & Sundberg J (2002). Spectral distribution of solo voice and accompaniment in pop music. *TFH-QPSR, KTH*, Vol 43: 31-35.
<http://www.speech.kth.se/qpsr/tmh/#V43>
- Öster A-M (2002). The relationship between residual hearing and speech intelligibility - Is there a measure that could predict a prelingually profoundly deaf child's possibility to develop intelligible speech? *TFH-QPSR, KTH*, Vol 43: 51-56. <http://www.speech.kth.se/qpsr/tmh/#V43>

* *Centrum för talteknologi (CTT); Centre for Speech Technology*