

Publications 2003



Dissertations

Bell L. *Linguistic adaptations in spoken human-computer dialogues. Empirical studies of user behavior.* Department of Speech, Music and Hearing, KTH, Stockholm, October 3, 2003. ISBN 91-7283-579-6.

Beskow J. *Talking Heads - Models and Applications for Multimodal Speech Synthesis.* Department of Speech, Music and Hearing, KTH, Stockholm. June 3, 2003.
<http://www.speech.kth.se/~beskow/thesis>. ISBN 91-7283-536-2.

Gobl C. *The voice source in speech communication. Production and perception experiments involving inverse filtering and synthesis.* Department of Speech, Music and Hearing, KTH, Stockholm. December 12, 2003. ISBN 91-7283-630-X.

Seward A. *Efficient methods for automatic speech recognition.* Department of Speech, Music and Hearing, KTH, Stockholm. December 17, 2003. ISBN 91-7283-657-1.

Thomason M. *From Air to Aria. Relevance of respiratory behaviour to voice function in classical western vocal art.* Department of Speech, Music and Hearing, KTH, Stockholm. May 16, 2003. ISBN 91-628-5703-7.

Granqvist S (2003). *Computer methods for voice analysis.* Department of Speech, Music and Hearing, KTH, Stockholm. March 28, 2003. ISBN 91-7283-461-7.

Licentiate thesis

Dahl S. *Striking movements: movement strategies and expression in percussive playing.* Department of Speech, Music and Hearing, KTH, Stockholm. May 9, 2003. ISBN 9-7283-480-3.

Journal Publications

Bretos J, Sundberg J (2003). Measurements of vibrato parameters in long sustained crescendo notes as sung by ten sopranos. *J Voice* 17: 343-353.

Engwall O (2003). Combining MRI, EMA & EPG measurements in a three-dimensional tongue model. *Speech Comm* 41: 303-329.

Goebel W & Bresin R (2003). Measurement and reproduction accuracy of computer-controlled grand pianos. *J Acoust Soc Am* 114(4): 2273-2283.

Granqvist S (2003). The self-to-other ratio applied as a phonation detector for voice accumulation. *Log Phon Vocol* 28: 71-80.

Granqvist S (2003). The Visual Sort and Rate method for perceptual evaluation in listening tests. *Log Phon Vocol* 28: 109-116.

Granqvist S & Hammarberg B (2003). The correlogram: a visual display of periodicity. *J Acoust Soc Am* 114 (5) 2934-2945.

Granqvist S, Hertegård S, Larsson H & Sundberg J (2003). Simultaneous analysis of vocal fold vibration and transglottal airflow. Exploring a new experimental set-up. *J Voice* 17/3: 319-330.

Heldner M (2003). On the reliability of overall intensity and spectral emphasis as acoustic correlates of focal accents in Swedish. *J Phonetics* 31(1): 39-62.

Hincks R (2003). Speech technologies for pronunciation feedback and evaluation. *ReCall* 15(1): 3-2.

Lindström E, Juslin N, Bresin R & Williamon A (2003). Expressivity comes from within your soul: A questionnaire study of music students' perspectives on expressivity. *Research Studies in Music Education*, 20: 23-47.

Maniscalco M, Weitzberg E, Sundberg J, Sofia M & Lundberg JO (2003). Assessment of nasal and sinus nitric oxide output using single-breath humming exhalations. *Eur Resp J*, 22: 323-329.

Rocchesso D, Bresin R & Fernström M (2003). Sounding object, *IEEE Multimedia Magazine* 10/2: 42-52.

Stone RE Jr, Cleveland TF, Sundberg J & Prokop (2003). Aerodynamic and acoustical measures of speech, operatic and Broadway vocal styles in a professional female singer. *J Voice* 17/3: 283-297.

Sundberg J, Friberg A & Bresin R (2003). Attempts to reproduce a pianist's expressive timing with Director Musices performance rules. *J New Music Res.* 22: 317-325.

Szabo A, Hammarberg B, Granqvist S & Södersten M (2003). Methods to study pre-school teachers' voice at work – Simultaneous recordings with a voice accumulator and a DAT recorder. *Log Phon Vocol* 28: 29-39.

Conference Proceedings

Allwood J & Cerrato L (2003). A study of gestural feedback expressions. *The First Nordic Symposium on Multimodal Communication*, Copenhagen, September 23-24, 2003 (CTT*).

Bell L & Gustafson J (2003). Child and adult speaker adaptation during error resolution in a publicly available spoken dialogue system. *Proc of EuroSpeech 2003*, Geneva, Switzerland, Sept 1-4, 613-616 (CTT).

Bell L, Gustafson J & Heldner M (2003). Prosodic adaptation in human-computer interaction. *Proc of ICPhS, XV Intl Conference of Phonetic Sciences*, Barcelona, Spain, Aug 3-9, 2003. 2453-2456 (CTT).

Beskow J, Engwall O & Granström B(2003). Simultaneous measurements of facial and intraoral articulation. *Proc of Fonetik 2003*, Umeå University, Dept of Philosophy and Linguistics PHONUM 9: 57-60 (CTT).

Blomberg M & Elenius D (2003). Collection and recognition of children s speech in the PF-Star project. *Proc of Fonetik 200,3 Umeå*

University, Dept of Philosophy and Linguistics PHONUM 9: 81-84 (CTT).

Bresin R & Juslin P (2003). Real-time visualization of musical expression. *Proc of ESCOM 5*, Hanover, Germany, Sept 8-10, 2003.

Bresin R, Dahl S, Rath M, Marshall M & Moynihan B (2003). Controlling the virtual bodhran - the vodhran. In: Bresin R (Ed.) *Proc of SMAC 2003, Stockholm Music Acoustics Conference*, Aug 6-9, 2003. 2: 685-688.

Bresin R, Hansen KF & Dahl S (2003). The Radio Baton as configurable musical instrument and controller. In: Bresin R (Ed.) *Proc of SMAC 2003, Stockholm Music Acoustics Conference*, Aug 6-9, 2003. 2: 689—691.

Canazza S, Friberg A, Rodà A & Zanon P (2003). Expressive Director: a system for the real-time control of music performance synthesis. In: Bresin R (Ed.) *Proc of SMAC 2003, Stockholm Music Acoustics Conference*, Aug 6-9, 2003. II: 521-524.

Carlson R & Swerts M (2003). Perceptually based prediction of upcoming prosodic breaks in spontaneous Swedish speech materials. *Proc of ICPhS, XV Intl Conference of Phonetic Sciences*, Barcelona, Spain, Aug 3-9, 2003. 79-82 (CTT).

Carlson R & Swerts M (2003). Relating perceptual judgments of upcoming prosodic breaks to F0 features. *Proc of Fonetik 2003*, Umeå University, Dept of Philosophy and Linguistics PHONUM 9, 181-184 (CTT)

Cerrato L & D'Imperio M (2003). Duration and tonal characteristics of short expressions in Italian. *Proc of ICPhS, XV Intl Conference of Phonetic Sciences*, Barcelona, Spain, Aug 3-9, 2003. (CTT).

Cerrato L & Skhiri M (2003). A method for the analysis and measurement of communicative head movements in human dialogues. *Proc of AVSP 2003, ISCA Tutorial and Research Workshop on Audio Visual Speech Processing*, St Jorioz, France, Sept 4-7, 2003. 251-256 (CTT).

Dahl S & Friberg A (2003). What can the body movements reveal about a musician's emotional intention? In: Bresin R (Ed.). *Proc of SMAC 03, Stockholm Music Acoustics Conference*, Aug 6-9, 2003. II: 599-602.

* Center for Speech Technology

- Dahl S & Friberg A (2003). Expressiveness of musician's body movements in performances on marimba. *Gesture Workshop 2003. The 5th Intl Workshop on Gesture and Sign Language Based Human-Computer Interaction*, Genoa, April 15-17, 2003.
- Dahl S & Granqvist S (2003). Looking at perception of continuous tempo drift - a new method for estimating internal drift and Just Noticeable Difference. In: Bresin R (Ed.), *Proc of SMAC03, Stockholm Music Acoustics Conference*, Aug 6-9, 2003. II: 595-598.
- Engwall O & Beskow J (2003). Resynthesis of 3D tongue movements from facial data. *Proc EuroSpeech 2003*, Geneva, Switzerland, Aug 3-9, 2003. 2261-2264 (CTT).
- Engwall O (2003). A revisit to the application of MRI to the analysis of speech production. Testing our assumptions. *6th Intl Seminar on Speech Production*, Sydney, Dec 7-10, 2003 (CTT).
- Engwall O & Beskow J (2003). The effect of corpus choice on statistical articulatory modeling. *6th Intl Seminar on Speech Production*, Sydney, Dec 7-10, 2003 (CTT).
- Fant G, Kruckenberg A & Barbosa-Ferreira J (2003). Individual variations in pausing. A study of read speech. *Proc of Fonetik 2003, PHONUM 9*, Umeå University: 193-196.
- Fontana F & Bresin R (2003). Physics-based sound synthesis and control: crushing, walking and running by crumpling sounds. *XIV Colloquium on Musical Informatics (XIV CIM 2003)*, Firenze, May 8-10, 2003.
- Galembo A (2003). Quality of treble piano tones. In: Bresin R (Ed.) *Proc of SMAC03, Stockholm Music Acoustics Conference*, Aug 6-9, 2003. I: 147-150.
- Galembo A & Askenfelt A (2003). Phase randomisation in piano bass tones. In: Bresin R (Ed.) *Proc of SMAC03, Stockholm Music Acoustics Conference*, Aug 6-9, 2003. I: 151-154.
- Galembo A & Askenfelt A (2003). Quality assessment of musical instruments - effects of multimodality, *ESCOM5, 5th Triennial Conference of the European Society for the Cognitive Sciences of Music*, Hannover, Sept 8-13, 2003.
- Goebel W & Bresin R (2003). Measurement and reproduction accuracy of computer-controlled grand pianos. In: Bresin R (Ed.) *Proc of SMAC 03, Stockholm Music Acoustics Conference*, Aug 6-9, 2003. I: 155-158.
- Goebel W, Bresin R & Galembo A (2003). The piano action as the performer's interface: Timing properties, dynamic behaviour and the performer's possibilities. In: Bresin R (Ed.) *Proc of SMAC 03, Stockholm Music Acoustics Conference*, Aug 6-9, 2003. I: 159-162.
- Granström B & House D (2003). Multimodality and speech technology: Verbal and non-verbal communication in talking agents. *Proc of EuroSpeech 2003*, Geneva, Switzerland, Sept 1-4, 2901-2904 (CTT).
- Guettler K, Schoonderwaldt E & Askenfelt A (2003). Bow speed or position – which one influences spectrum the most? In: Bresin R (Ed.) *Proc of SMAC 03, Stockholm Music Acoustics Conference*, Aug 6-9, 2003. 67-70.
- Hansen KF & Bresin R (2003). DJ scratching performance techniques: Analysis and synthesis. In: Bresin R (Ed.) *Proc of SMAC 03, Stockholm Music Acoustics Conference*, Aug 6-9, 2003. II: 693-696.
- Heldner M & Megyesi B (2003). Exploring the prosody-syntax interface in conversations. *Proc 15th ICPhS, XV Intl Conference of Phonetic Sciences*. Barcelona, Spain, Aug 3-9, 2003. 2501-2504 (CTT).
- Heldner M & Megyesi B (2003). The acoustic and morpho-syntactic context of prosodic boundaries in dialogs. *Proc of Fonetik 2003, Umeå University, Dept of Philosophy and Linguistics PHONUM 9*: 117-120 (CTT).
- Hertegård S, Larsson H & Granqvist S (2003). Vocal fold resonances at low and high pitch tuning - preliminary results. In: Bresin R (Ed.) *Proc of SMAC 03, Stockholm Music Acoustics Conference*, Aug 6-9, 2003. (Poster).
- Hincks R (2003). Pronouncing the academic word list: Features of L2 student oral presentations. *Proc of ICPhS, XV Intl Conference of Phonetic Sciences*. Barcelona, Spain, Aug 3-9, 2003. 1545-1549 (CTT).
- Hincks R (2003). Tutors, tools and assistants for the L2 user. *Proc of Fonetik 2003, Umeå University, Dept of Philosophy and Linguistics PHONUM 9*: 173-176 (CTT).

- House D (2003). Hesitation and interrogative Swedish intonation. *Proc of Fonetik 2003, Umeå University, Dept of Philosophy and Linguistics PHONUM 9*: 185-188 (CTT).
- House D (2003). Perceiving question intonation: the role of pre-focal pause and delayed focal peak. *Proc of ICPHS, XV Intl Conference of Phonetic Sciences*. Barcelona, Spain, Aug 3-9, 2003. 755-758 (CTT).
- House D (2003). Perception of tone with particular reference to tonal alignment. In: Kaji S (Ed.) *Proc of the International Symposium on Crosslinguistic Studies of Tonal Phenomena 2002*. Tokyo: Institute for Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies. Dec 17-19, 2002.
- Jande P-A (2003). Evaluating rules for phonological reduction in Swedish. *Proc of Fonetik 2003, Umeå University, Dept of Philosophy and Linguistics, PHONUM 9*: 149-152 (CTT).
- Jande P-A (2003). Phonological reduction in Swedish. *Proc of ICPHS, XV Intl Conference of Phonetic Sciences*. Barcelona, Spain, Aug 3-9, 2003. 2557-2560 (CTT).
- Karlsson I (2003). The SYNFACE project - a status report. *Proc of Fonetik 2003, Umeå University, Dept of Philosophy and Linguistics PHONUM 9*: 61-64 (CTT).
- Karlsson I, Faulkner A & Salvi G (2003). SYNFACE - a talking face telephone *Proc of EuroSpeech 2003*, Geneva, Switzerland, Sept 1-4, 1297-1300 (CTT).
- Magnuson T & Hunnicutt S (2003). Support for the construction of sentences and phrases for symbol users. *Proc of Fonetik 2003, Umeå University, Dept of Philosophy and Linguistics PHONUM 9*: 85-88 (CTT).
- Mathews MV, Friberg A, Bennett G, Sapp C & Sundberg J (2003). A marriage of the Director Musices program and the conductor program. In: Bresin R (Ed.). *Proc of SMAC 03, Stockholm Music Acoustics Conference*, Aug 6-9, 2003. I: 13-16.
- Minematsu N, Kurata G & Hirose K (2003). Corpus-based analysis of production and perception of Japanese English in view of the entire phonemic system of English. In: Solé MJ, Recasens D & Romero J (eds). *Proc of ICPHS 2003, XV Intl Conference of Phonetic Sciences*, Barcelona, Spain, Aug 3-9, 2003. 1569-1572.
- Minematsu N, Guo C & Hirose K (2003). CART-based factor analysis of intelligibility reduction in Japanese English. *Proc of EuroSpeech 2003*, Geneva, Switzerland, Sept 1-4, 2069-2072.
- Minematsu N (2003). Phonetic tree analysis and its application to automatic estimation of the segmental intelligibility. *Proc of EuroSpeech 2003*, Geneva, Switzerland, Sept 1-4.
- Minematsu N, Yamauchi K & Hirose K (2003). Automatic estimation of perceptual age using speaker modeling techniques. *Proc of EuroSpeech 2003*, Geneva, Switzerland, Sept 1-4. 3005-3008.
- Minematsu N, Osaki K & Hirose K (2003). Improvement of non-native speech recognition by effectively modeling frequently observed pronunciation habits. *Proc of EuroSpeech 2003*, Geneva, Switzerland, Sept 1-4. 2597-2600.
- Nordstrand M, Svanfeldt G, Granström B & House D (2003). Measurements of articulatory variation and communicative signals in expressive speech. *Proc of AVSP'03, ISCA Tutorial and Research Workshop on Audio Visual Speech Processing*, St Jorioz, France, Sept 4-7, 2003. 233-238 (CTT).
- Pakucs B (2003). Towards dynamic multi-domain dialogue processing. *Proc of EuroSpeech 2003*, Geneva, Switzerland, Sept 1-4, 741-744 (CTT).
- Pakucs B (2003). SesAME: A framework for personalized and adaptive speech interfaces. *Proc of the EACL-03 Workshop on Dialogue Systems: interaction, adaptation and styles of management*. Budapest. pp 95-102. (CTT)
- Rinman M-L, Friberg A, Kjellmo I, Camurri A, Cirotteau D, Dahl S, Mazzarino B, Bendiksen B & McCarthy H (2003). EPS - an interactive collaborative game using non-verbal communication. In: Bresin R (Ed.). *Proc of SMAC 03, of the Stockholm Music Acoustics Conference*, Aug 6-9, 2003. II: 561-563.
- Salvi G (2003). Accent clustering in Swedish using the Bhattacharyya distance. *Proc of ICPHS, XV Intl Conference of Phonetic*

Sciences. Barcelona, Spain, Aug 3-9, 2003. 1149-1152 (CTT).

Salvi G (2003). Using accent information in ASR models for Swedish. *Proc of EuroSpeech 2003*, Geneva, Switzerland, Sept 1-4, 2677-2680 (CTT).

Schoonderwaldt, E., Guettler, K., Askenfelt, A. (2003). Effect of the width of the bow hair on the violin string spectrum. In: Bresin R (Ed.). *Proc of SMAC 03, Stockholm Music Acoustics Conference*, Aug 6-9, 2003. 91-94.

Seward A (2003). Low-latency incremental speech transcription in the Synface project. *Proc of EuroSpeech 2003*, Geneva, Switzerland, Sept 1-4, 1141-1144 (CTT).

Siciliano C, Williams G, Beskow J & Faulkner A (2003). Evaluation of a Multilingual Synthetic Talking Face as a communication Aid for the Hearing Impaired. *Proc of ICPhS, XV Intl Conference of Phonetic Sciences*. Barcelona, Spain, Aug 3-9, 2003. pp 131-134. (CTT).

Sjölander K (2003). An ensemble method for the alignment of sound and its transcription. In: Bresin R (Ed.). *Proc of SMAC 03, Stockholm Music Acoustics Conference*, Aug 6-9, 2003. II: 743-746. (CTT)

Sjölander K (2003). An HMM-based system for automatic segmentation and alignment of speech. *Proc of Fonetik 2003, Umeå University, Dept of Philosophy and Linguistics PHONUM 9*: 93-96 (CTT).

Sjölander P (2003). Perceptual relevance of the 5 kHz spectral region to sex identification in children's singing voices, In: Bresin R (Ed.). *Proc of SMAC 03, Stockholm Music Acoustics Conference*, Aug 6-9, 2003. 503-506.

Skantze D & Dahlbäck N (2003). Auditory icons support for navigation in speech-only interfaces for room-based design metaphors. *Proc of ICAD 2003, Intl Community for Auditory Display*, Boston University, USA, July 6-9, 2003. 140-143 (CTT).

Skantze G (2003). Exploring human error handling strategies: implications for spoken dialogue systems. *Proc of ISCA Tutorial and Research Workshop on Error Handling in Spoken Dialogue Systems*, 71-76 (CTT).

Svanfeldt G, Nordstrand M, Granström B & House D (2003). Measurements of articulatory

variation in expressive speech. *Proc of Fonetik 2003, Umeå University, Dept of Philosophy and Linguistics, PHONUM 9*: 53-56 (CTT).

Ternström S, Bohman M & Södersten S (2003). Very loud speech over simulated environmental noise tends to have a spectral peak in the F1 region. 145th ASA meeting, Nashville TN, April 2003. *J Acoust Soc Am* 113 (4/2): 2296 (Abstr).

Ternström S (2003). Sex and the singer: gender categorization aspects of singing voice. 145th ASA meeting, Nashville TN, April 2003. *J Acoust Soc Am* 113 (4/2): 2242 (Abstr).

Öster A-M, House D, Hatzis A & Green P (2003). Testing a new method for training fricatives using visual maps in the Ortho-Logo-Paedia project (OLP). *Proc of Fonetik 2003, Umeå University, Dept of Philosophy and Linguistics PHONUM 9*: 89-92 (CTT).

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

Granqvist S (2003). The self-to-other ratio applied as a phonation detector for voice accumulation. *TMH-QPSR, KTH*, 45: 23-33.

Granqvist S, Hertegård S, Larsson H & Sundberg J (2003). Simultaneous analysis of vocal fold vibration and transglottal airflow; Exploring a new experimental set-up. *TMH-QPSR, KTH*, 45: 35-46.

Nienkerke-Springer A, McAllister A & Sundberg J (2003). Effects of family therapy on children's voices. *TMH-QPSR, KTH*, 45: 75-85.

Nordenberg M & Sundberg J (2003). Effect on LTAS of vocal loudness variation. *TMH-QPSR, KTH*, 45: 93-100.

Sundberg J (2003). Research on the singing voice in retrospect. *TMH-QPSR, KTH*, 45: 11-22.

Sundberg J, Friberg A & Bresin R (2003). Musician's and computer's tone inter-onset interval in Mozart's Piano Sonata K 332, 2nd mvt, bar 1-20, *TMH-QPSR, KTH*, 45: 47-59.

Thomasson M (2003). Belly-in or belly-out? Effects of inhalatory behaviour and lung

volume on voice function in male opera singers. *TMH-QPSR, KTH*, 45: 61-73.

Thomasson M (2003). Effects of lung volume on the glottal voice source and the vertical laryngeal position in male professional operatic singers. *TMH-QPSR, KTH*, 45: 1-9.

Zangger Borch D, Sundberg J, Lindestad PA & Thalén M (2003). Vocal fold vibration and voice source aperiodicity in phonatorily distorted singing. *TMH-QPSR, KTH*, 45: 87-91.

Books, Book Chapters

Arlinger S, Brorsson B, Lagerbring C, Leijon A, Rosenhall U & Schersten T (2003).

Hörapparat för vuxna. Nytt och kostnader (Hearing Aids for Adults. Benefits and Costs). Stockholm: *Statens Beredning för Medicinsk Utvärdering*, SBU. Nr 164.

Bresin R & Dahl S (2003). Experiments on gestures: walking, running, and hitting. In:

Rocchesso D & Fontana F (Eds). *The Sounding Object*, Mondo Estremo, Florence, Italy, pp. 111-136.

Bresin R, Hansen KF, Dahl S, Rath M, Marshall M & Moynihan B (2003). Devices for manipulation and control of sounding objects: The Vodhran and the Invisiball. In: Rocchesso D & Fontana F (Eds). *The Sounding Object*, Mondo Estremo, Florence, Italy, pp. 271 -296.

Cerrato L & Paoloni A (2003). Utilizzo dei parametri della fonetica acustica nell'identificazione del parlante in ambito forense. In: Cosi P, Magno Caldognetto E, Zamboni A (eds). *Voce, Canto Parla to, Studi in Onore di Franco Ferrero*. Unipress 2003, 59-66.

Hansen KF & Bresin R (2003). Complex gestural audio control: the case of scratching. In: Rocchesso D & Fontana F (Eds). *The Sounding Object*. Mondo Estremo, Florence, Italy, 221-269.