

Curriculum vitae – Torbjørn Svendsen

Key qualifications:

Research and teaching experience in machine learning and artificial intelligence for over 40 years, specifically in speech technology. Speech recognition and synthesis are areas which require multidisciplinary collaboration, both with regards to applications and to the development of systems and methods, and we have engaged in such collaborations since 1990.

I have extensive project management experience through a series of research projects, and have also administrative experience as Department Head and Faculty Vice Dean. I have also a number of appointments on the national level, e.g. as chair of the program board of the VERDIKT ICT research program in the Research Council of Norway.

I have been full professor at NTNU since 1995. In the period 2015-2021 I was also Director of NTNU Digital, which is NTNU's strategic initiative for research and application of digital technologies. In that capacity I was heavily involved with the establishment of the Norwegian Open AI Lab (NAIL). I was also involved in the planning and programming of the Code of Ethics workshop in 2019, addressing ethical issues of digitalization and AI.

I have a long-standing and extensive international engagement, including research visits at leading institutions such as AT&T Bell Labs, AT&T Shannon Labs and MIT. I have been Vice President of the International Speech Communication Association (ISCA), the world's largest organization for researchers in spoken language processing.

I have >100 peer reviewed publications, mainly on various aspects of speech technology. I have been the supervisor of 19 PhD graduates and ~100 Master's degree students.

PERSONAL INFORMATION

Family name, First name:	Svendsen, Torbjørn		
Date of birth:	23.05.1955	Sex:	Male
Nationality:	Norwegian		

*HIGHER EDUCATION/OTHER TRAINING

	Subjects/degree/	Name of institution, country
1985	Dr.ing.	NTH, Norway
1981	Pedagogical education for engineers	NTH, Norway
1979	Siv.ing. (MSc)	NTH, Norway

POSITIONS (academic, business, industry, public sector, national or international organisations)

Current Position

	Job title/name of employer/country
1988 -	Professor, NTNU, Norway

Previous positions held (list)

	Job title/name of employer/country
2015-2021	Director, NTNU Digital, Norway
1983-1984	Doctoral scholarship, NTH (on leave from SINTEF) , Norway
1981-1987	Scientist, SINTEF, Norway
1979-1981	Teaching assistant, NTH, Norway

PROJECT MANAGEMENT EXPERIENCE (if applicable)

	Project/topic/role in project/funding from
2014-2020	AULUS/Speech technology/Project manager/Forskningsrådet
2006-2010	SIRKUS/Speech recognition/Project manager/Forskningsrådet
2003-2007	VOCALS/ Speech recognition/Project manager/Forskningsrådet
2003-2008	FONEMA/Speech synthesis/Project manager/Forskningsrådet
2004-2008	Rundkast/Speech databases/Project manager/NTNU
1997-2001	SPODIS/Spoken dialogue systems/Project manager/Forskningsrådet

Relevant public committees/appointments

- ☐ Member, appropriation committee for the ICT program IKTPLUSS in the Research Council of Norway (2015-2019)
- ☐ Program board chair for the ICT research program VERDIKT in the Research Council of Norway (2005-2010)
- ☐ Member of ICT Forum, a group mandated with proposing a research strategy for ICT for the Research Council of Norway (2000-2001); the program board of the ICT-programme in the Research Council of Norway (2001-2006), and the User Forum for the program "IT for the Disabled" (2000-2003, also in the Research Council of Norway)
- ☐ Chair of a working group formed by the Norwegian Language Council assigned with proposing an "Action Plan for Norwegian Language and ICT" (2001-2002).
- ☐ Chair of a project group appointed by the Ministry of Culture and Church Affairs for "Consolidating and Increasing the Availability of Norwegian Human Language Resources" (2002)
- ☐ Member of advisory board, Norwegian Language Bank ("Språkbanken", 2010-2017)

- Member of expert group appointed by the Norwegian Technology Council (2022)
- Expert group member, digitalization in UNIT (Norwegian Directorate for ICT and joint services in higher education and research)
- Expert group member contributing to the Norwegian national strategy Digital21

Academic memberships and roles

- Elected member, Norwegian Academy of Technological Sciences
- Life Senior Member, IEEE
- Vice President (currently Secretary) International Speech Communication Association (ISCA)

International community

- Reviewer for several international scientific journals, including IEEE Trans. on Communications, IEEE Trans. on Signal Processing (previously Acoustics, Speech and Signal Processing); IEEE Trans. on Speech, Audio and Language Processing; IEEE Trans. on Multimedia; EURASIP Journal on Applied Signal Processing; and Speech Communications
- Member of *Speech Communication* journal Editorial Board
- Reviewer for various conferences and workshops on speech and signal processing.
- Interspeech: Organizing and Technical committees 2000, Area Chair 2012, 2018-2021, regular reviewer
- IEEE: Member of IEEE Signal Processing Society's Speech Technical Committee 1998-2001, IEEE Flanagan Award committee 2011-2014, regular reviewer for ICASSP
- Reviewer for EU's Language Engineering program and the Information Society Research Programme of the Academy of Finland.
- Project reviews for the Norwegian, Australian, Swiss, Dutch, Belgian and South African Research Councils
- Opponent/member of examination boards for 23 doctoral theses
- COST Actions: Vice chair, COST action 278 "Spoken Language Interaction in Telecommunication". Working group chair, COST actions 232 "Speech Recognition over the telephone line" and 249 "Continuous speech recognition over the telephone"
- EU framework programmes: Steering committee member and researcher, ESPRIT SAM, Advisory Scientific Board member ACORNS, partner MSCN Sound-to-sense and e-Ladda
- Board member, Nordic Graduate School of Language Technology

Research visits:

Apr 1986 - Apr 1987	Speech Research Dept., AT&T Bell Labs, Murray Hill, NJ, USA
Jun 1991 - Aug 1991	Speech Research Dept., AT&T Bell Labs, Murray Hill, NJ, USA
Jul 1996 - Jun 1997	Griffith University, Brisbane, Australia
Jun 2000 – Sep 2000	Speech and Image Research Dept, AT&T Labs, Florham Park, NJ, USA
Jul 2002 – Jun 2003	Queensland University of Technology, Brisbane, Australia
Jan 2013 - Jul 2013	MIT Computer Science and Artificial Intelligence Lab, Cambridge, MA, USA
Aug 2022 - Dec 2022	Delft University of Technology, Delft, The Netherlands
Jan 2023 - Jun 2023	Univ. of Enna Kore, Enna, Italy

**Internal appointments**

Department Head, Dept. of Telecommunications, 1992-1995

Vice Dean, Faculty of Electrical Engineering and Telecommunications(1999 - 2001)

Group Head, Signal Processing (several periods from 1992-2012)

A list of selected publications:

- T. Svendsen, F. K. Soong:** "On the Automatic Segmentation of Speech Signals", Proc. ICASSP'87.
- T. Svendsen, K. K. Paliwal, E. Harborg, P. O. Husøy:** "An Improved Sub-Word Based Speech Recognizer", Proc. ICASSP'89
- P.O. Husøy, T. Svendsen:** "ANN-based speech recognition using a preprocessor for non-linear time compression", Proc. Eurospeech'91, Genova, Italy
- A.Basu, T.Svendsen:** "A Time-Frequency Segmental Neural Network for Phoneme Recognition", Proc. ICASSP'93, pp. 1509-1512, Minneapolis, April 1993
- T.Svendsen, F.K.Soong, H.Purnhagen:** "Optimizing baseforms for HMM-based speech recognition", Proc. Eurospeech'95, Madrid, Sept.1995
- E.Harborg, T.Holter, M.H.Johnsen, T.Svendsen:** "On-line captioning of TV-programs for the hearing impaired", Proc. Eurospeech'99, Budapest, Sept. 1999
- T.Holter, T.Svendsen:** "Maximum likelihood modelling of pronunciation variation", Speech Communications, Nov. 1999
- M.H.Johnsen, T.Svendsen, T.Amble, T.Holter, E.Harborg:** "TABOR - A Norwegian Spoken Dialogue System for Bus Travel Information", Proc. ICSLP'2000, Beijing, Oct. 2000.
- T.A. Myrvoll, K. K. Paliwal, T. Svendsen:** "Fast Adaptation using Constrained Affine Transformations with Hierarchical Priors", Proc. Eurospeech 2001, Aalborg, september 2001
- T. Svendsen:** "Roles for Speech Technology in the Information Society", in *E. Panzar (ed.): Perspectives on the Age of the Information Society*, Reports of the Information Research Programme of the Academy of Finland, pp. 229-238, Tampere 2002
- E.Wong, T.Martin, T.Svendsen, S.Sridharan:** "Multilingual Phone Clustering for Recognition of Spontaneous Indonesian Speech Utilising Pronunciation Modelling Techniques", Proc. Eurospeech 2003, Geneva, Sept. 2003
- T.Martin, T.Svendsen, S.Sridharan:** "Cross-Lingual Pronunciation Modelling for Indonesian Speech Recognition", Proc. Eurospeech 2003, Geneva, Sept. 2003
- I.Amdal, T.Svendsen:** "Unit Selection Synthesis Database Development Using Utterance Verification", Proc. Eurospeech 2005, Lisbon 2005
- T.Skogstad, T.Svendsen:** "Distributed ASR Using Speech Coder Data for Efficient Feature Vector Representation", Proc. Eurospeech 2005, Lisbon 2005
- D.Meen, T.Svendsen:** "Improving Phone Label Alignment by Utilizing Voicing Information", Proc. SPECOM-2005, Patras, Greece, Oct. 2005
- S.M.Siniscalchi, T.Svendsen, C.-H.Lee:** "Towards Bottom-Up Phone Recognition" Proc. ASRU 2007, Kyoto, Japan, Dec. 2007
- S.M.Siniscalchi, T.Svendsen, C.-H.Lee:** "Toward A Detector-Based Universal Phone Recognizer ", ICASSP 2008, Las Vegas, Nevada, April 2008
- I.Amdal, O.M.Strand, J.Almberg, T.Svendsen:** "RUNDKAST: An Annotated Norwegian Broadcast News Corpus", LREC 2008, Marrakech, Morocco, May 2008
- S.M.Siniscalchi, T.Svendsen, C.-H.Lee:** "A Phonetic Feature Based Lattice Rescoring Approach to LVCSR", Proc. ICASSP 2009, Taipei, Taiwan, April 2009
- S.M.Siniscalchi, J.Reed, T.Svendsen, C.-H.Lee:** "Exploring Universal Attribute Characterization of Spoken Languages for Spoken Language Recognition", Proc. Interspeech 2009, Brighton, Sept. 2009
- T.Mertens, D.Schneider, A.B.Næss, T.Svendsen:** "Lexicon Adaptation for Subword Speech Recognition", Proc. ASRU 2009, Merano, Italy, Dec. 2009
- L.Adde, B.Reveil, J.-P.Martens, T.Svendsen:** "A Minimum Classification Error approach to pronunciation variation modeling of non-native proper names", Proc. Interspeech 2010; Makuhari, Japan, Sept 2010
- D.Meen, T.Svendsen:** "The NTNU Concatenative Speech Synthesizer", in Proc. Blizzard 2010 Workshop, Kyoto, Japan, Sept. 2010
- T.Skogstad, T.Svendsen:** "Frequency-Warped and Stabilized Time-Varying Cepstral Coefficients", Proc. Interspeech, Florence, Sept. 2011

- M.Soufifar, M.Kockmann, L.Burget, O.Plhot, O.Glembek, T.Svendsen:** "iVector Approach to Phonotactic Language Recognition", Proc. Interspeech, Florence, Sept. 2011
- S.M. Siniscalchi, T. Svendsen, C.-H. Lee:** "A Bottom-Up Modular Search Approach to Large Vocabulary Continuous Speech Recognition", IEEE Transactions on Audio, Speech, and Language Processing, April 2013
- S.M.Siniscalchi, J. Reed, T.Svendsen, C.-H. Lee:** "Universal Attribute Characterization of Spoken Languages for Automatic Spoken Language Recognition", Computer Speech and Language, Jan. 2013
- J.B. Hamar, R.S. Doddipatla, T. Svendsen:** "Non-Negative Duration HMM", Proc. IEEE International Workshop on Machine Learning for Signal Processing, Southampton, Sept. 2013
- R.S. Doddipatla, T. Svendsen:** "Synthetic Speaker Models Using VTLN to Improve the Performance of Children in Mismatched Speaker Conditions for ASR", Proc. Interspeech, Lyon, Sept. 2013
- S. M. Siniscalchi, T. Svendsen, C.-H. Lee:** "An Artificial Neural Network Approach to Automatic Speech Processing", Neurocomputing 140, pp326-338, Sept. 2014, <http://dx.doi.org/10.1016/j.neucom.2014.03.005>
- T.Svendsen, J.B.Hamar:** "Combining NdHMM and Phonetic Feature Detection for Speech Recognition", Proc. EUSIPCO 2015, Nice, Sept. 2015
- A.S.Shahrehabaki, N.Olfati, A.S.Imran, S.M.Siniscalchi, T.Svendsen:** "A phonetic-level analysis of different input features for articulatory inversion", Proc. Interspeech2019, Graz, Sept. 2019
- A.S.Shahrehabaki, N.Olfati, S.M.Siniscalchi, G.Salvi, T.Svendsen:** "Transfer learning of articulatory information through phone information", Proc. Interspeech2020, Shanghai, Oct. 2020
- A.S.Shahrehabaki, S.M.Siniscalchi, G.Salvi, T.Svendsen:** "Sequence-to-sequence articulatory inversion through time convolution of sub-band frequency signals", Proc. Interspeech2020, Shanghai, Oct. 2020
- A.S. Shahrehabaki, G. Salvi, T. Svendsen, S.M. Siniscalchi:** "Acoustic-to-Articulatory Mapping With Joint Optimization of Deep Speech Enhancement and Articulatory Inversion Models". IEEE/ACM Transactions on Audio, Speech, and Language Processing (TASLP), vol 30, pp. 135-147, 2022, doi:10.1109/TASLP.2021.3133218
- J. Rugayan, T. Svendsen and G. Salvi,** "Semantically Meaningful Metrics for Norwegian ASR Systems", Proc. Interspeech 2022, doi: 10.21437/Interspeech.2022-817

Full Publication List

- [1]: **T.Svendsen:** "Effects of Rayleigh Fading on Medium Bandwidth Speech Coders", Proc. EUSIPCO'83, pp 539-542
- [2]: **T.Svendsen:** "Tree Encoding of the LPC Residual", Proc.I CASSP'84
- [3]: **T.Svendsen, K.K.Paliwal:** "Comparative Evaluations of LP Analysis Methods for Multi-Pulse Excited LP Coding of Speech", Proc. of The Institute of Acoustics Autumn Conference, Windermere, Great Britain, 1984.
- [4]: **K.K.Paliwal, T.Svendsen:** "A Study of Three Coders (Sub-Band, RELP and MPE) for speech with Additive White Noise", Proc. ICASSP'85.
- [5]: **T.Svendsen:** "Multi-Dimensional Quantization Applied to Predictive Coding of Speech", Ph.D thesis, The Norwegian Institute of Technology, March 1985.
- [6]: **T. Svendsen:** "Speech processing in Norway", Proc. Nordic Conf. on Speech Processing, Trondheim, March 1986.
- [7]: **M. H. Johnsen, T. Svendsen:** "A Speaker Independent Isolated Word Recognizer for a Norwegian Vocabulary", Proc. Nordic Conf. on Speech Processing, Trondheim, March 1986
- [8]: **T.Svendsen:** "Multi-Dimensional Quantization Applied to Predictive Coding of Speech", Proc. ICASSP'86.
- [9]: **T. Svendsen, F. K. Soong:** "On the Automatic Segmentation of Speech Signals", Proc. ICASSP'87.
- [10]: **T. Svendsen:** "Speech Recognition- An Overview", Telecommunications, Vol. 21, No. 12, Dec. 1987.

- [11]: **T. Svendsen, K. K. Paliwal, E. Harborg, P. O. Husøy:** "Experiments with a Sub-Word Based Speech Recognizer", 2nd Australian Conf. on Speech Science and Technology, Sydney, Australia, 1988.
- [12]: **T. Svendsen, K. K. Paliwal, E. Harborg, P. O. Husøy:** "An Improved Sub-Word Based Speech Recognizer", Proc. ICASSP'89
- [13]: **T. Svendsen:** "Speech Processing Activities in the Nordic Countries", Proc. Speech Tech'90 .
- [14]: **T. Svendsen, K. K. Paliwal, E. Harborg, P. O. Husøy:** "An Improved Sub-Word Based Speech Recognizer", Proc. ICASSP'89
- [15]: **P.O. Husøy, T. Svendsen:** "ANN-based speech recognition using a preprocessor for non-linear time compression", Proc. Eurospeech'91, Genova, Italy
- [16]: **B.Ribbum, A.Perkis, T.Svendsen:** "The self-excited, stochastic multipulse coders", Journal of Electrical and Electronics Engineering, Australia, pp. 183-187, September 1991
- [17]: **A.Paoloni, T.Svendsen, B.Kaspar, D.Johnston, G.Hult:** "COST232: Speech Recognition Over the Telephone Line", Proc. EUROSPEECH'93, pp. 1845-1848, Berlin, Sept. 1993
- [18]: **A.Basu, T.Svendsen:** "A Time-Frequency Segmental Neural Network for Phoneme Recognition", Proc. ICASSP'93, pp. 1509-1512, Minneapolis, April 1993
- [19]: **T.Svendsen:** "Efficient Quantization of Speech Spectral Information", Proc. EUROSPEECH'93, pp. 1143-1146, Berlin, Sept. 1993
- [20]: **T.Svendsen:** "Segmental Quantization of Speech Spectral Information", Proc. ICASSP'94, Adelaide, April, 1994
- [21]: **T.Svendsen, B.T.Jensen:** "Access Control Using Speaker Verification", Proc. NORSIG-94, pp. 130-133, Ålesund, June, 1994
- [22]: **T.Holter, T.Svendsen:** "Training of lexica for subword-based speech recognisers", Proc. Norsig-95, pp. 57-61, Stavanger, 1995
- [23]: **T.Svendsen, F.K.Soong, H.Purnhagen:** "Optimizing baseforms for HMM-based speech recognition", Proc. Eurospeech'95, Madrid, Sept.1995
- [24]: **T.Holter, T.Svendsen:** "Combined optimisation of baseforms and subword models for an HMM based speech recognizer", Proc. ISSPA'96, pp. 321--324, Gold Coast, Australia, July, 1996
- [25]: **J.Pihl, T.Svendsen, M.H.Johnsen:** "A VLSI implementation of PDF computations in HMM based speech recognition", Proc. IEEE TENCON, Perth, Australia, Nov. 1996
- [26]: **T.Holter, T.Svendsen:** "A comparison of lexicon-building methods for subword-based speech recognisers", Proc. IEEE TENCON, Perth, Australia, Nov. 1996
- [27]: **T.Holter, T.Svendsen:** "A joint segmentation and labelling scheme for use in acoustic subword based speech recognition". Proc. Norsig'97, pp. 30-33, Tromsø, May 1997
- [28]: **T.Holter, T.Svendsen:** "Incorporating Linguistic Knowledge and automatic baseform generation in acoustic subword unit based speech recognition", Proc. Eurospeech'97, Sept. 1997
- [29]: **T.Holter, T.Svendsen:** "Combined Optimisation of Baseforms and Model Parameters in Speech Recognition Based on Acoustic Subword Units", Proc. IEEE Workshop on Automatic Speech Recognition and Understanding, Santa Barbara, Dec. 1997
- [30]: **T.Holter, T.Svendsen:** "Maximum likelihood modelling of pronunciation variation", ESCA Tutorial and Research Workshop on "Modeling pronunciation variation for ASR", Kerkrade, The Netherlands, May 1998
- [31]: **T. Svendsen:** "Blir norsk gresk for taleteknologien" ("Will Norwegian be Greek for Speech Technology" - in Norwegian), Språknytt, No. 4, 1998
- [32]: **I.Amdal, T.Holter, T.Svendsen:** "Maximum likelihood pronunciation modelling of Norwegian natural numbers for automatic speech recognition", Proc. Norsig'99, Asker, Sept. 1999
- [33]: **T. Svendsen:** "Taleteknologi" ("Speech technology" - in Norwegian), in *Språk i Norden*, B. Lindgren et. al, editors, pp. 9-19, Novus forlag, 1999
- [34]: **E.Harborg, T.Holter, M.H.Johnsen, T.Svendsen:** "On-line captioning of TV-programs for the hearing impaired", Proc. Eurospeech'99, Budapest, Sept. 1999

- [35]: **E.Harborg, T.Holter, M.H.Johnsen, T.Svendsen:** "Generation of closed captions for live TV-programs using speech recognition", Proc. Norsig'99, pp. 141-145, Asker, Sept. 1999.
- [36]: **I.Amdal, T.Holter, T.Svendsen:** "Maximum likelihood pronunciation modelling of Norwegian natural numbers for automatic speech recognition", Proc. Norsig'99, pp. 146-150, Asker, Sept. 1999.
- [37]: **T.Holter, T.Svendsen:** "Maximum likelihood modelling of pronunciation variation", Speech Communications, Nov. 1999
- [38]: **E.Harborg, T.Holter, M.H.Johnsen, T.Svendsen:** "Subtitling of live broadcast TV-programs for the hearing impaired", Proc. AAATE'99, http://www.fernuni-hagen.de/FTB/aaate99/paper/99_103/99_103.htm, Dusseldorf, Nov. 1999
- [39]: **M.H.Johnsen, T.Holter, T.Svendsen, T.Holter:** "Stochastic Modelling of Semantic Content for Use in a Spoken Dialogue System", Proc. ICSLP'2000, Beijing, Oct. 2000.
- [40]: **T.Holter, M.H.Johnsen, E.Harborg, T.Svendsen:** "ASR-Based Subtitling of Live TV-Programs for the Hearing Impaired", Proc. ICSLP'2000, Beijing, Oct. 2000.
- [41]: **M.H.Johnsen, T.Svendsen, T.Amble, T.Holter, E.Harborg:** "TABOR - A Norwegian Spoken Dialogue System for Bus Travel Information", Proc. ICSLP'2000, Beijing, Oct. 2000.
- [42]: **I.Amdal, T.Holter, T.Svendsen:** "Modellering av uttalevariasjon for automatisk talegjenkjenning", Nordlyd, vol. 28-2000, pp. 74-87, 2000
- [43]: **T.A. Myrvoll, K. K. Paliwal, T. Svendsen:** "Fast Adaptation using Constrained Affine Transformations with Hierarchical Priors", Proc. Eurospeech 2001, Aalborg, september 2001
- [44]: **T.Svendsen:** "Nordisk forskningssamarbeid innen språkteknologi", Språknytt, vol 3/2001, pp. 24-28, Sept. 2001
- [45]: **M.H.Johnsen, E.Harborg, T.Svendsen, T.Amble, T.Holter, T.A.Myrvoll, T.Norgård:** "SPODIS – Spoken Dialog Systems for Telephony", Proc. Norsig 2001, Trondheim, October 2001
- [46]: **T. Svendsen:** "Roles for Speech Technology in the Information Society", in *E. Panzar (ed.): Perspectives on the Age of the Information Society*, Reports of the Information Research Programme of the Academy of Finland, pp. 229-238, Tampere 2002
- [47]: **I.Amdal, T.Svendsen:** "Evaluation of Pronunciation Variants in the ASR Lexicon for Different Speaking Styles", Proc. LREC-2002, Las Palmas, May 2002
- [48]: **E.Wong, T.Martin, T.Svendsen, S.Sridharan:** "Multilingual Phone Clustering for Recognition of Spontaneous Indonesian Speech Utilising Pronunciation Modelling Techniques", Proc. Eurospeech 2003, Geneva, Sept. 2003
- [49]: **T.Martin, T.Svendsen, S.Sridharan:** "Cross-Lingual Pronunciation Modelling for Indonesian Speech Recognition", Proc. Eurospeech 2003, Geneva, Sept. 2003
- [50]: **T.Svendsen:** "Speech Technology: Past, Present and Future", Teletronikk, vol 99, no. 2, pp. 6-18, 2003
- [51]: **T.Svendsen:** "Pronunciation Modeling for Speech Technology", Proc. 2004 International Conference on Signal Processing and Communications; Bangalore, 2004
- [52]: **I.Bjørkan, T.Svendsen, S.Farner:** "Comparing Spectral Distance Measures for Join Cost Optimization in Concatenative Speech Synthesis", Proc. Eurospeech 2005, Lisbon 2005
- [53]: **I.Amdal, T.Svendsen:** "Unit Selection Synthesis Database Development Using Utterance Verification", Proc. Eurospeech 2005, Lisbon 2005
- [54]: **T.Skogstad, T.Svendsen:** "Distributed ASR Using Speech Coder Data for Efficient Feature Vector Representation", Proc. Eurospeech 2005, Lisbon 2005
- [55]: **D.Meen, T.Svendsen:** "Improving Phone Label Alignment by Utilizing Voicing Information", Proc. SPECOM-2005, Patras, Greece, Oct. 2005
- [56]: **I.Amdal, T.Svendsen:** "FonDat1: A Speech Synthesis Corpus for Norwegian", Proc. LREC-2006, Genova, Italy, May 2006
- [57]: **I.Amdal, T.Svendsen:** "Log Likelihood Ratio Based Annotation Verification of a Norwegian Speech Synthesis Database", Proc. NORSIG 2006, Reykjavik, June 2006

- [58]: **T.Svendsen**: "Articulatory Features And Segmental Information for Automatic Speech Recognition", ESF Exploratory Workshop on Models of Language Evolution, Acquisition and Processing, Leuven, Belgium, Nov. 2007
- [59]: **S.M.Siniscalchi, T.Svendsen, C.-H.Lee**: "Towards Bottom-Up Phone Recognition", Proc. ASRU 2007, Kyoto, Japan, Dec. 2007
- [60]: **S.M.Siniscalchi, T.Svendsen, C.-H.Lee**: "Toward A Detector-Based Universal Phone Recognizer ", ICASSP 2008, Las Vegas, Nevada, April 2008
- [61]: **I.Amdal, O.M.Strand, J.Almberg, T.Svendsen**: "RUNDKAST: An Annotated Norwegian Broadcast News Corpus", LREC 2008, Marrakech, Morocco, May 2008
- [62]: **T.Skogstad, T.Svendsen**: "Time-Varying Cepstral Coefficients", ISCA ITRW on Speech Analysis and Processing for Knowledge Discovery, Aalborg, Denmark, June 2008
- [63]: **S.M.Siniscalchi, Ø.Birkenes, M.H.Johnsen, T.Svendsen**: "Joint Optimization of Event Detectors and Evidence Merger for Continuous Speech Recognition", ISCA ITRW on Speech Analysis and Processing for Knowledge Discovery, Aalborg, Denmark, June 2008
- [64]: **S.M.Siniscalchi, T.Svendsen, C.-H.Lee**: "A Penalized Logistic Regression Approach to Detection Based Phone Classification", Proc. Interspeech 2008, Brisbane, Australia, Sept. 2008
- [65]: **S.M.Siniscalchi, T.Svendsen, C.-H.Lee**: "A Phonetic Feature Based Lattice Rescoring Approach to LVCSR", Proc. ICASSP 2009, Taipei, Taiwan, April 2009
- [66]: **S.M.Siniscalchi, J.Reed, T.Svendsen, C.-H.Lee**: "Exploring Universal Attribute Characterization of Spoken Languages for Spoken Language Recognition", Proc. Interspeech 2009, Brighton, Sept. 2009
- [67]: **T.Mertens, D.Schneider, A.B.Næss, T.Svendsen**: "Lexicon Adaptation for Subword Speech Recognition", Proc. ASRU 2009, Merano, Italy, Dec. 2009
- [68]: **S.M.Siniscalchi, T.Svendsen, F.Sorbello, C.-H.Lee**: "Experimental Studies on Continuous Speech Recognition Using Neural Architectures with 'Adaptive' Hidden Activation Functions", Proc. ICASSP 2010, Dallas, March 2010
- [69]: **R.O.Sikveland, A.Oettl, I.Amdal, M.Ernestus, T.Svendsen, J.Edlund**: "SponTal-N: A Corpus of Interactional Spoken Norwegian", Proc. LREC 2010, Malta, May 2010
- [70]: **L.Adde, T.Svendsen**: "NameDat: A Database of English Proper Names Spoken by Native Norwegians", Proc. LREC 2010, Malta, May 2010
- [71]: **T.Skogstad, T.Svendsen**: "Intra-Frame Variability As a Predictor of Frame Classifiability", Proc. Interspeech 2010; Makuhari, Japan, Sept 2010
- [72]: **S.M. Siniscalchi1, J. Reed, T. Svendsen3, and C.-H Lee**: Exploiting Context-Dependency and Acoustic Resolution of Universal Speech Attribute Models in Spoken Language Recognition", Proc. Interspeech 2010; Makuhari, Japan, Sept 2010
- [73]: **L.Adde, B.Reveil, J.-P.Martens, T.Svendsen**: "A Minimum Classification Error approach to pronunciation variation modeling of non-native proper names", Proc. Interspeech 2010; Makuhari, Japan, Sept 2010
- [74]: **D.Meen, T.Svendsen**: "The NTNU Concatenative Speech Synthesizer", in Proc. Blizzard 2010 Workshop, Kyoto, Japan, Sept. 2010
- [75]: **L. Adde and T. Svendsen**, "On the use of discriminative and non-discriminative pronunciation priors in pronunciation variation modeling of non-native proper names," 2010 IEEE Spoken Language Technology Workshop, San Francisco, Dec. 2010, pp. 229-234, doi: <http://10.1109/SLT.2010.5700856>, <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5700856>
- [76]: **L.Adde, T.Svendsen**: "Pronunciation Variation Modeling of Non-Native Proper Names by Discriminative Tree Search", Proc. IEEE ICASSP, Prague, May 2011, doi: <http://doi.org/10.1109/ICASSP.2011.5947461> , <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5947461>
- [77]: **S.M.Siniscalchi, T.Svendsen, C.-H.Lee**: "A Bottom-Up Stepwise Knowledge-Integration Approach to Large Vocabulary Continuous Speech Recognition Using Weighted Finite State Machines", Proc. Interspeech, Florence, Sept. 2011

- [78]: **T.Skogstad, T.Svendsen:** "Frequency-Warped and Stabilized Time-Varying Cepstral Coefficients", Proc. Interspeech, Florence, Sept. 2011
- [79]: **M.Souffar, M.Kockmann, L.Burget, O.Plhot, O.Glembek, T.Svendsen:** "iVector Approach to Phonotactic Language Recognition", Proc. Interspeech, Florence, Sept. 2011
- [80]: **Rodriguez-Fuentes, J. et al:** "Multi-Site Heterogeneous System Fusions For The Albayzin 2010 Language Recognition Evaluation", Proc. IEEE ASRU, Honolulu, Dec 2011
- [81]: **S.M.Siniscalchi, D.-H.Lyu, T.Svendsen, C.-H. Lee:** "Experiments on Cross-Language Attribute Detection and Phone Recognition With Minimal Target-Specific Training Data", IEEE Trans. on Audio, Speech and Language Processing, March 2012, doi: 10.1109/TASL.2011.2167610, <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6016213>
- [82]: **S.M. Siniscalchi, T. Svendsen, C.-H. Lee:** "A Bottom-Up Modular Search Approach to Large Vocabulary Continuous Speech Recognition", IEEE Transactions on Audio, Speech, and Language Processing, April 2013, doi: <https://doi.org/10.1109/TASL.2012.2234115> , <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6384711>
- [83]: **S.M.Siniscalchi, J. Reed, T.Svendsen, C.-H. Lee:** "Universal Attribute Characterization of Spoken Languages for Automatic Spoken Language Recognition", Computer Speech and Language, Jan. 2013, <https://doi.org/10.1016/j.csl.2012.05.001> , <https://reader.elsevier.com/reader/sd/pii/S0885230812000307?token=B8E6B113687AAD857AFC48468233AF26471DC918D347E83CC6FC83D8AF1D54383615D90F2A8D7F59EEF282CB28E5F296&originRegion=eu-west-1&originCreation=20221101085707>
- [84]: **J.B. Hamar, R.S. Doddipatla, T. Svendsen:** "Non-Negative Duration HMM", Proc. IEEE International Workshop on Machine Learning for Signal Processing, Southampton, Sept. 2013, doi: <https://doi.org/10.1109/MLSP.2013.6661976>, <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6661976>
- [85]: **R.S. Doddipatla, T. Svendsen:** "Synthetic Speaker Models Using VTLN to Improve the Performance of Children in Mismatched Speaker Conditions for ASR", Proc. Interspeech, Lyon, Sept. 2013, <http://dx.doi.org/10.21437/Interspeech.2013-743>, https://www.isca-speech.org/archive/pdfs/interspeech_2013/sanand13_interspeech.pdf
- [86]: **S. M. Siniscalchi, T. Svendsen, C.-H. Lee:** "An Artificial Neural Network Approach to Automatic Speech Processing", Neurocomputing 140, pp326-338, Sept. 2014, <http://dx.doi.org/10.1016/j.neucom.2014.03.005>
- [87]: **T.Svendsen, J.B.Hamar:** "Combining NdHMM and Phonetic Feature Detection for Speech Recognition", Proc. EUSIPCO 2015, Nice, Sept. 2015
- [88]: **A.S.Shahrebabaki, A.S.Imran, N.Olfati, T.Svendsen:** "Acoustic Feature Comparison for Different Speaking Rates", *Human-Computer Interaction. Interaction Technologies*. Springer 2018, pp. 176-189, ISBN 978-3-319-91250-9
- [89]: **A.S.Shahrebabaki, A.S.Imran, N.Olfati, T.Svendsen:** "A Comparative Study of Deep Learning Techniques on Frame-Level Speech Data Classification", *Circuits, systems, and signal processing* , vol 38, pp 3501-3520, 2019, <https://doi.org/10.1007/s00034-019-01130-0>
- [90]: **A.S.Imran, A.S.Shahrebabaki, N.Olfati, T.Svendsen:** "A Study on the Performance Evaluation of Machine Learning Models for Phoneme Classification", *ICMLC '19 Proceedings of the 2019 11th International Conference on Machine Learning and Computing*. Association for Computing Machinery (ACM), pp. 52-58, 2019. ISBN 978-1-4503-6600-7.
- [91]: **A.S.Imran, Z. Kastrati, T.Svendsen, A.Kurti:** "Text-Independent Speaker ID for Automatic Video Lecture Classification Using Deep Learning", *ICCAI '19 Proceedings of the 2019 5th International Conference on Computing and Artificial Intelligence*. Association for Computing Machinery (ACM), pp 175-180, 2019, ISBN 978-1-4503-6106-4.
- [92]: **A.S.Imran, V.Haflan, A.S.Shahrebabaki, N.Olfati, T.Svendsen:** "Evaluating Acoustic Feature Maps in 2D-CNN for Speaker Identification", *ICMLC '19 Proceedings of the 2019 11th International Conference on Machine Learning and Computing*. Association for Computing Machinery (ACM), pp. 211-216, 2019 ISBN 978-1-4503-6600-7.

- [93]: **A.S.Shahrehabaki, N.Olfati, A.S.Imran, S.M.Siniscalchi, T.Svendsen:** "A phonetic-level analysis of different input features for articulatory inversion", Proc. Interspeech2019, Graz, Sept. 2019, doi: 10.21437/Interspeech.2019-2526, https://www.isca-speech.org/archive/pdfs/interspeech_2019/shahrehabaki19_interspeech.pdf
- [94]: **A.S.Shahrehabaki, N.Olfati, S.M.Siniscalchi, G.Salvi, T.Svendsen:** "Transfer learning of articulatory information through phone information", Proc. Interspeech2020, Shanghai, Oct. 2020, doi: 10.21437/Interspeech.2020-1139, https://www.isca-speech.org/archive/pdfs/interspeech_2020/shahrehabaki20_interspeech.pdf
- [95]: **A.S.Shahrehabaki, S.M.Siniscalchi, G.Salvi, T.Svendsen:** "Sequence-to-sequence articulatory inversion through time convolution of sub-band frequency signals", Proc. Interspeech2020, Shanghai, Oct. 2020, doi: 10.21437/Interspeech.2020-1140, https://www.isca-speech.org/archive/pdfs/interspeech_2020/shahrehabaki20b_interspeech.pdf
- [96]: **A.S. Shahrehabaki, N. Olfati, A.S. Imran, M.H. Johnsen, S.M. Siniscalchi, T. Svendsen:** "A Two-Stage Deep Modeling Approach to Articulatory Inversion", IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2021, pp. 6453-6457, doi: 10.1109/ICASSP39728.2021.9413742, <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9413742>
- [97]: **A. S. Shahrehabaki, S. M. Siniscalchi, G. Salvi and T. Svendsen:** "A DNN Based Speech Enhancement Approach to Noise Robust Acoustic-to-Articulatory Inversion," 2021 IEEE International Symposium on Circuits and Systems (ISCAS), 2021, pp. 1-5, doi: 10.1109/ISCAS51556.2021.9401290.
- [98]: **A.S. Shahrehabaki, S.M. Siniscalchi, T. Svendsen,** "Raw Speech-to-Articulatory Inversion by Temporal Filtering and Decimation". Proc. Interspeech 2021, 1184-1188, doi: 10.21437/Interspeech.2021-1429, doi: 10.21437/Interspeech.2021-1429, https://www.isca-speech.org/archive/pdfs/interspeech_2021/shahrehabaki21_interspeech.pdf
- [99]: **A.S. Shahrehabaki, G. Salvi, T. Svendsen, S.M. Siniscalchi,** "Acoustic-to-Articulatory Mapping With Joint Optimization of Deep Speech Enhancement and Articulatory Inversion Models". *IEEE/ACM Transactions on Audio, Speech, and Language Processing (TASLP)*, vol 30, pp. 135-147, 2022, doi:10.1109/TASLP.2021.3133218, <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9640504>
- [100]: **Y. Getman, R. Al-Ghezi, K. Voskoboinik, T. Grósz, M. Kurimo, G. Salvi, T. Svendsen, S. Strömbergsson,** "wav2vec2-based Speech Rating System for Children with Speech Sound Disorder", Proc. Interspeech 2022, doi: 10.21437/Interspeech.2022-10103, https://www.isca-speech.org/archive/pdfs/interspeech_2022/getman22_interspeech.pdf
- [101]: **J. Rugayan, T. Svendsen and G. Salvi,** "Semantically Meaningful Metrics for Norwegian ASR Systems", Proc. Interspeech 2022, doi: 10.21437/Interspeech.2022-817, https://www.isca-speech.org/archive/pdfs/interspeech_2022/rugayan22_interspeech.pdf
- [102]: **Z. Fan, X. Cao, G. Salvi, T. Svendsen:** "Using Modified Adult Speech as Data Augmentation for Child Speech Recognition", ICASSP 2023
- [103]: **P.E. Solberg, P. Ortiz, G. Salvi, T. Svendsen, P. Parsons:** "Improving Generalization of Norwegian ASR with Limited Linguistic Resources", NoDaLiDa 2023
- [104]: **P. Parsons, K. Kvale, T. Svendsen, G. Salvi:** "A Character Based Analysis of Impacts of Dialects on End-to-end Norwegian ASR", NoDaLiDa 2023
- [105]: **J. Rugayan,, G. Salvi, T. Svendsen:** "Perceptual and Task-Oriented Assessment of a Semantic Metric for ASR Evaluation", Proc. Interspeech 2023, Dublin, Aug. 2023
- [106]: **X. Cao, Z. Fan, T. Svendsen, G. Salvi:** "An Analysis of Goodness of Pronunciation for Child Speech", Proc. Interspeech 2023, Dublin, Aug. 2023.
- [107]: **Y. Getman, N. Phan, R. Al-Ghezi, E. Voskoboinik, M. Singh, T. Grósz, M. Kurimo, G. Salvi, T. Svendsen, S. Strömbergsson, A. Smolander, S. Ylinen.:** "Developing an AI-assisted Low-resource Spoken Language Learning App for Children," in IEEE Access, doi: 10.1109/ACCESS.2023.3304274.
- [108]: **F. B. Gelderblom, T.V. Tronstad, T. Svendsen, T.A. Myrvoll:** "On the Predictive Power of Objective Intelligibility Metrics for the Subjective Performance of Deep Complex Convolutional Recurrent

