
Jonathan Harrington

Institut für Phonetik und Sprachverarbeitung, Ludwig-Maximilians-Universität München

Human speech processing and its relationship to historical sound change



In human speech processing, the task is to understand how words and their speech sounds are transmitted from a speaker to a hearer. Historical sound change very occasionally happens because perception involves actively reconstructing and interpreting the speech signal which creates a leaky transmission of information from the speaker to the hearer. This talk reviews recent advances in applying technology from speech tools and machine learning for explaining how spoken interaction between individuals can over short and long time scales cause sound change in the entire population. Using experimental techniques to model

sound change can, in turn, provide fresh insights into explaining how information flows from speech signals to words and their constituent speech sounds in human speech communication.

Hanna Fischer

Forschungszentrum Deutscher Sprachatlas, Philipps-Universität Marburg

Regionalsprachliche Tonaufnahmen als Ressource für die elektronische Sprachsignalverarbeitung



Für eine Optimierung von Spracherkennung, aber auch für die automatisierte Auswertung von Sprachaufnahmen müssen die Analysemodelle und -verfahren an natürlicher, alltäglicher Sprache trainiert werden. Diese ist in allen Regionen des deutschsprachigen Raums regional geprägt. Mit den digitalisierten Tonkorpora des Forschungszentrums Deutscher Sprachatlas (DSA) stehen umfangreiche Bestände von Sprachaufnahmen der deutschen Dialekte und Regionalakzente aus unterschiedlichen Provenienzen und Erhebungssettings zur Verfügung. Die Aufnahmen

eignen sich hervorragend für die Weiterentwicklung der einschlägigen Technologien und werden am DSA in unterschiedlichen Forschungsprojekten (u. a. zur KI-basierten Analyse der Dynamik dialektalen Sprechens sowie zur Erschließung raumbildender phonotaktischer Muster im Deutschen) eingesetzt. Der Beitrag informiert über die verfügbaren Sprachdaten sowie über die aktuellen Forschungsaktivitäten am DSA.

Inhaltsverzeichnis – Contents

Sprachsynthese – Speech Synthesis	1
<i>Bernd J. Kröger</i>	
Articulatory Speech Synthesis in the Context of Speech Research and Speech Technology: Review and Prospect	1
<i>Antoine Serrurier</i>	
Can Deep Learning Help to Understand Speech Production Mechanisms?	9
Interaktion & Dialog – Interaction & Dialogue	17
<i>Anna-Maria Meck</i>	
How May I Interrupt? Linguistic Design Guidelines for Proactive In-Car Voice Assistants	17
<i>Mika Rebensburg, Stefan Hillmann, Nils Feldhus</i>	
Automatic User Experience Evaluation of Goal-Oriented Dialogs Using Pre-Trained Language Models	25
<i>Philine Görzig, Jan Nehring, Stefan Hillmann, Sebastian Möller</i>	
A Comparison of Module Selection Strategies for Modular Dialog Systems	33
<i>Stefan Hillmann, Philine Görzig, Sebastian Möller</i>	
Automatic Generation of Website-Based Multi-Turn Question-Answering Dialog Systems	41
Visualisierung und Verarbeitung – Visualisation and Processing	49
<i>Christian Kleiner, Peter Birkholz</i>	
Comparison of Object Tracking Algorithms for Larynx Phantom Move- ments in Ultrasound Videos	49
<i>Frederick Kukla, Vanessa Reichel</i>	
Anwendung des MFCC-Plotters zur Erfassung cepstraler Unterschiede in emotionaler Sprache	57
<i>Christoph Draxler</i>	
Analysis of Transcriptions – A Pilot Study	65
Sprachpathologie – Speech Pathology	67
<i>Ian S. Howard, Julian McGlashan, Adrian J. Fourcin</i>	
Training a CNN to Estimate Voice Pathology from Connected Speech Us- ing EGG to Automatically Label the Dataset for Voicing	67
<i>Konstantin Sering, Paul Schmidt-Barbo</i>	
Somatosensory Feedback in PAULE	75
<i>Eugenia Rykova, Mathias Walther</i>	
Concept for Semantic Error Analysis in a Mobile Application for Speech and Language Therapy Support	83

<i>Hans-Günter Hirsch, Christian Neumann, Yannic Tiggelkamp, Riccardo Fiorista, Stefan Knecht, Alfons Schnitzler, Katja Biermann-Ruben, Dietmar Bothe, Günter Bleimann, Hendrike Frieg</i>	RehaLingo – Towards a Speech Training System for Aphasia	90
Phonetik – Phonetics		98
<i>Jürgen Trouvain, Raphael Werner</i>	Muster der Sprechatmung in verschiedenen Sprechstilen – Eine Pilotstudie	98
<i>Alexis Deighton MacIntyre, Raphael Werner</i>	An Automatic Method for Speech Breathing Annotation	106
<i>Simon Betz</i>	Hesitation Lengthening Elicitation and Detection via Target Words in a Card Game Study	114
<i>Birgitte Poulsen, Ocke-Schwen Bohn, Christoph Draxler</i>	A First Report on a Perceptual Training Study Using Percy	122
Automatische Spracherkennung – Automatic Speech Recognition		130
<i>Dominik Wagner, Sebastian P. Bayerl, Tobias Bocklet</i>	Implementing Easy-to-Use Recipes for the Switchboard Benchmark . . .	130
<i>Aaricia Herygers, Vass Verkhodanova, Matt Coler, Odette Scharenborg, Munir Georges</i>	Bias in Flemish Automatic Speech Recognition	138
Poster		146
<i>Neda Mousavi, Sven Grawunder</i>	Persian Speaker Classification Using Rhythmic Features	146
<i>Stefan Ciba, Mohammed Krini, Amir Rajabi</i>	Approach to Speaker-Generalized Spectral Envelope Estimation by Deep Recurrent Neural Network for Speech Reconstruction in a Speech Enhancement System	154
<i>Robert Wardenga, Daniel Vogel, Felix Gräßer, Mira Schielke, Leonard Starke, Rico Petrick, Torsten Rex, Jens Lehmann</i>	iDOKS: Ein integriertes Dokumentationssystem zur Zusammenfassung von Gesprächen und Meetings	161
<i>Jan Nehring, Nils Feldhus, Akhyar Ahmed</i>	Adapters for Resource-Efficient Deployment of NLU Models	169
<i>Matthias Busch, Malte Kania, Tom Assmann, Ingo Siegert</i>	Radlogistik als Anwendungsgebiet für Digitale Sprachassistenten – Ein Diskussionsbeitrag	177
<i>Harald Höge</i>	Synchrony of θ - Oscillations in Speech Perception and Speech Production	185
Show & Tell		190
<i>Felix Burkhardt</i>	Nkululeko: A Template Based System for Fast Machine Learning Experiments on Speaker Characteristics	190

<i>Thomas Ranzenberger, Tobias Bocklet, Steffen Freisinger, Lia Frischholz, Munir Georges, Kevin Glocker, Aaricia Herygers, René Peinl, Korbinian Riedhammer, Fabian Schneider, Christopher Simic, Khabbab Zakaria</i>	
The Hochschul-Assistenz-System HanS: an ML-Based Learning Experience Platform	192
<i>Christoph Draxler, Julian Pömp</i>	
Transcription Portal – A Zero-configuration Workbench for Transcribing Spoken Language Recordings	194
Kinderstimmen – Child Speech	195
<i>Christopher Gebauer, Lars Rumberg, Jörn Ostermann</i>	
Pronunciation Modelling for Children’s Speech	195
<i>Hanna Ehlert, Edith Beaulac, Maren Wallbaum, Christopher Gebauer, Lars Rumberg, Jörn Ostermann, Ulrike Lüdtk</i>	
Collecting and Annotating Natural Child Speech Data – Challenges and Interdisciplinary Perspectives	203
Emotion	210
<i>Felix Burkhardt, Uwe Reichel, Florian Eyben, Björn Schuller</i>	
Going Retro: Astonishingly Simple Yet Effective Rule-Based Prosody Modelling for Speech Synthesis Simulating Emotion Dimensions	210
<i>Jan Hintz, Andreas Wendemuth, Ingo Siegert</i>	
Cross-Reliability Benchmark Test for Preserving Emotional Content in Speech–Synthesis Related Datasets	218