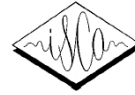




UZH
alumni



XVII AISV CONFERENCE

**Speaker Individuality in Phonetics and Speech Sciences:
Speech Technology and Forensic Applications**

Thursday 4th - Friday 5th February 2021



Book of abstracts

XVII AISV Conference

Associazione Italiana Scienze della Voce
Thursday 4th - Friday 5th February 2021

Hosted by University of Zurich
(online)

Organising Committee

Stephan Schmid (chair),
Camilla Bernardasci, Volker Dellwo,
Dalila Dipino, Davide Garassino, Michele Loporcaro,
Stefano Negrinelli, Elisa Pellegrino,
Dieter Studer-Joho

Student Assistant

Seraina Nadig

Scientific Committee

CINZIA AVESANI, ISTC-CNR, Padova
PIER MARCO BERTINETTO, Scuola Normale Superiore di Pisa
SILVIA CALAMAI, Università di Siena
FRANCESCO CANGEMI, Universität zu Köln
CHIARA CELATA, Università degli Studi di Urbino Carlo Bo
SONIA CENCESCHI, Scuola universitaria professionale della Svizzera italiana
FRANCESCO CUTUGNO, Università degli Studi di Napoli Federico II
VOLKER DELLWO, Universität Zürich
ANNA DE MEO, Università degli Studi di Napoli L'Orientale
LORENZO FILIPPONIO, Humboldt-Universität zu Berlin
HELEN FRASER, University of New England
PETER FRENCH, University of York
VINCENZO GALATÀ, ISTC-CNR, Padova
DAVIDE GARASSINO, Universität Zürich
BARBARA GILI FIVELA, Università del Salento
MIRKO GRIMALDI, Università del Salento
LEI HE, Universität Zürich
WILLEMJN HEEREN, Universiteit Leiden
MICHAEL JESSEN, Bundeskriminalamt, Wiesbaden
THAYABARAN KATHIRESAN, Universität Zürich
FELICITAS KLEBER, Ludwig-Maximilians-Universität München
MICHELE LOPORCARO, Universität Zürich
PAOLO MAIRANO, Université de Lille
GIOVANNA MAROTTA, Università di Pisa
PIETRO MATURI, Università degli Studi di Napoli Federico II
KIRSTY MCDUGALL, University of Cambridge
CHIARA MELUZZI, Università degli Studi di Pavia
FRANCIS NOLAN, University of Cambridge
ANTONIO ORIGLIA, Università degli Studi di Napoli Federico II
ELISA PELLEGRINO, Universität Zürich
MICHAEL PUCHER, Institut für Schallforschung, Wien
ANTONIO ROMANO, Università degli Studi di Torino
LUCIANO ROMITO, Università della Calabria
PIER LUIGI SALZA, Socio onorario AISV
CARLO SCHIRRU, Università degli Studi di Sassari
SANDRA SCHWAB, Universität Zürich; Université de Fribourg
MARIO VAYRA, Università di Bologna
ALESSANDRO VIETTI, Libera Università di Bolzano
CLAUDIO ZMARICH, ISTC-CNR, Padova

Table of contents

Plenary Lectures	1
HELEN FRASER	
Forensic transcription: Scientific and legal perspectives	2
KIRSTY MCDUGALL	
Ear-Catching versus Eye-Catching? Some Developments and Current Challenges in Earwitness Identification Evidence	3
General Session	4
NICOLAS AUDIBERT, CÉCILE FOUGERON AND ESTELLE CHARDENON	
Do you remain the same speaker over 21 recordings?	5
ANGELIKA BRAUN	
The quest for speaker individuality – a challenge for forensic phonetics	7
SILVIA CALAMAI, MARIA FRANCESCA STAMULI AND ALESSANDRO CASELLATO	
Un percorso condiviso per la redazione di un <i>Vademecum</i> sulla conservazione, la descrizione, l'uso e il riuso delle fonti orali	9
HONGLIN CAO AND XIAOLIN ZHANG	
The Current Situation of the Application of Evidence of Forensic Phonetics in Courts of China	11
LEONARDO CONTRERAS ROA, PAOLO MAIRANO, CAROLINE BOUZON AND MARC CAPLIEZ	
The acquisition of /s/ - /z/ in a phonemic vs neutralised context: comparing French _{L1} , Italian _{L1} and Spanish _{L1} learners of L2 English	13
SONIA D'APOLITO AND BARBARA GILI FIVELA	
Realizzazione di suoni nativi nel parlato di Italiano L2 da parte di parlanti francofoni: Interazione tra accuratezza e contesto	15
STEFON FLEGO AND JON FORREST	
Interspeaker variation in anticipatory coarticulation: A whole-formant approach	17

SALVATORE GIANNINÒ, CINZIA AVESANI, GIULIANO BOCCI AND MARIO VAYRA	
Prosodia implicita ed esplicita: convergenze e divergenze nella risoluzione di ambiguità sintattiche globali	19
ADRIANA HANULÍKOVÁ	
Do faces speak volumes? A life span perspective on social biases in speech comprehension and evaluation	21
LEI HE	
Characterizing speech rhythm using spectral coherence between jaw displacement and speech temporal envelope	23
THAYABARAN KATHIRESAN, ARJUN VERMA AND VOLKER DELLWO	
Gender bias in voice recognition: An i-vector-based gender-specific automatic speaker recognition study	25
KATHARINA KLUG, MICHAEL JESSEN AND ISOLDE WAGNER	
Collection and analysis of multi-condition audio recordings for forensic automatic speaker recognition	27
ADRIAN LEEMANN, PÉTER JESZENSZKY, CARINA STEINER AND HANNAH HEDEGARD	
Earwitness evidence accuracy revisited: Estimating age, weight, height, education, and geographical origin	29
ADAS LI, PETER FRENCH, VOLKER DELLWO AND ELEANOR CHODROFF	
Analysing the effect of language on speaker-specific speech rhythm in Cantonese-English bilinguals	32
JUSTIN LO	
Seeing the trees in the forest: Diagnosing individual performance in likelihood ratio based forensic voice comparison	34
ROSALBA NODARI AND SILVIA CALAMAI	
I silenzi dei matti. Gli spazi ‘vuoti’ del parlato nell’archivio sonoro di Anna Maria Bruzzone	36
BENJAMIN O'BRIEN, ALAIN GHIO, CORINNE FREDOUILLE, JEAN-FRANÇOIS BONASTRE AND CHRISTINE MEUNIER	
Discriminating speakers using perceptual clustering interface	38
HANNA RUCH, ANDREA FRÖHLICH AND MARTIN LORY	
Clustering of unknown voices	40

SIMONA SBRANNA, CATERINA VENTURA, AVIAD ALBERT AND MARTINE GRICE	
Prosodic marking of information status in L1 Italian and L2 German	42
LOREDANA SCETTINO, SIMON BETZ, FRANCESCO CUTUGNO AND PETRA WAGNER	
Hesitations and Individual Variability in Italian Tourist Guides' Speech	44
LAURA SMORENBURG AND WILLEMIJN HEEREN	
Forensic value of acoustic-phonetic features from Standard Dutch nasals and fricatives	46
BRUCE WANG, VINCENT HUGHES AND PAUL FOULKES	
System performance and speaker individuality in LR-based forensic voice comparison	48
<i>Poster Presentations</i>	50
ALICE ALBANESI, SONIA CENCESCHI, CHIARA MELUZZI AND ALESSANDRO TRIVILINI	
Italian monozygotic twins' speech: a preliminary forensic investigation	51
CHIARA BERTINI, PAOLA NICOLI, NICCOLÒ ALBERTINI AND CHIARA CELATA	
A 3D model of linguopalatal contact for VR biofeedback	53
SILVIA CALAMAI AND CECILIA VALENTINI	
Sull'insegnamento della pronuncia italiana negli anni sessanta a bambini e a stranieri	55
MEIKE DE BOER AND WILLEMIJN HEEREN	
Language-dependency of /m/ in L1 Dutch and L2 English	57
VALENTINA DE IACOVO, MARCO PALENA AND ANTONIO ROMANO	
La variazione prosodica in italiano: l'utilizzo di un chatbot Telegram per la didattica assistita per apprendenti di italiano L2 e nella valutazione linguistica delle conoscenze disciplinari	59
MARCO FARINELLA, MARCO CARNAROGLIO AND FABIO CIAN	
Una nuova idea di "impronta vocale" come strumento identificativo e riabilitativo	61

CHLOË FARR, GRACELLIA PURNOMO, AMANDA CARDOSO, ARIAN SHAMEI AND BRYAN GICK	
Speaker Accommodations and VUI Voices: Does Human-likeness of a Voice Matter?	63
MANUELA FRONTERA	
Radici identitarie e mantenimento linguistico. Il caso di un gruppo di <i>heritage speakers</i> di origine calabrese	65
DAVIDE GARASSINO, DALILA DIPINO AND FRANCESCO CANGEMI	
Modeling intonation in interaction. A new approach to the intonational analysis of questions in (semi-)spontaneous speech	67
GLENDIA GURRADO	
Sulla codifica e decodifica della sorpresa	69
LEI HE AND WILLEMIJN HEEREN	
Between-speaker variability in dynamic formant characteristics in spontaneous speech	71
ELLIOT HOLMES	
Using Phonetic Theory to Improve Automatic Speaker Recognition	73
ANNA HUSZÁR, VALÉRIA KREPSZ, ALEXANDRA MARKÓ AND TEKLA ETELKA GRÁCZI	
Formant variability in five Hungarian vowels with regard to speaker Discriminability	75
KATHARINA KLUG, CHRISTIN KIRCHHÜBEL, PAUL FOULKES AND PETER FRENCH	
How robust are perceptual and acoustic observations of breathiness to mobile phone transmission?	77
CAROLINA LINS MACHADO	
A cross-linguistic study of between-speaker variability in intensity dynamics in L1 and L2 spontaneous speech	79
MARCO MARINI, MAURO VIGANÒ, MASSIMO CORBO, MARINA ZETTIN, GLORIA SIMONCINI, BRUNO FATTORI, CLELIA D'ANNA, MASSIMILIANO DONATI AND LUCA FANUCCI	
The first Italian Dysarthric Speech Database for improving daily living of severely dysarthric people	81
ÁLVARO MOLINA-GARCÍA	
Acoustics and Perception do not match in Andalusian Spanish	83

UMAR MUHAMMAD, PETER FRENCH AND ELEANOR CHODROFF	
A Comparative Analysis of Nigerian Linguist Native Speakers and Untrained Native Speakers Categorising Four Accents of Nigerian English	86
ELISA PELLEGRINO AND VOLKER DELLWO	
Dynamics of short-term cross-dialectal accommodation. A study on Grison and Zurich German	88
ALEJANDRA PESANTEZ	
L2 speakers' individual differences in the acoustic properties of the front-high English vowels: The case of Ecuadorian speakers	90
DUCCIO PICCARDI AND FABIO ARDOLINO	
Variazione e <i>user engagement</i> . Un approfondimento sulla ludicizzazione dei protocolli d'inchiesta linguistica	92
CLAUDIA ROSWANDOWITZ, THAYABARAN KATHIRESAN, ELISA PELLEGRINO, VOLKER DELLWO AND SASCHA FRÜHHOLZ	
First indications for speaker individuality and speech intelligibility in state-of-the-art artificial voices	94
YU ZHANG, LEI HE, KARNTHIDA KERDPOL AND VOLKER DELLWO	
Between-speaker variability in intensity slopes: The case of Thai	96
CLAUDIO ZMARICH, SERENA BONIFACIO, MARIA GRAZIA BUSÀ, BENEDETTA COLAVOLPE, MARIAVITTORIA GAIOTTO AND FRANCESCO OLIVUCCI	
Coarticulation and VOT in four Italian children from 18 to 48 months of age	98
<i>Satellite Workshop</i>	100
MICHAEL JESSEN	
Workshop on automatic and semiautomatic speaker recognition	101
<i>Round table</i>	102
Current trends and issues in forensic phonetics research	103

Between-speaker variability in dynamic formant characteristics in spontaneous speech

Lei He¹ and Willemijn Heeren²

¹ *Department of Computational Linguistics, University of Zürich*

² *Leiden University Centre for Linguistics, Leiden University*

lei.he@uzh.ch; w.f.l.heeren@hum.leidenuniv.nl

Introduction

The temporal characteristics of speech articulation have received relatively little attention in forensic phonetics, because directly characterizing speaker-specific articulatory movements is almost impossible; kinematic data of articulators are absent from case materials. However, forensic speech scientists may instead focus on acoustic properties in the speech signal that are – although not entirely – modulated by the articulatory movements. For example, Dellwo and colleagues measured speech rhythm in terms of the durational variability of various phonetic intervals (e.g., Dellwo et al. 2015, Leemann et al. 2014) or syllabic intensity variability (e.g., He and Dellwo 2014, 2016); McDougall (2006) approached formant trajectories using least-squares polynomial approximations; and He and Dellwo (2017) measured the dynamic characteristics of intensity contours. Their study found that measures based on the speeds of intensity decreases (i.e., negative intensity dynamics) explained approximately 70% of between-speaker variability, pointing to a possibility that the mouth-closing gestures may contain more speaker-specific information.

More recently, He et al. (2019) combined the ideas of both McDougall (2006) and He and Dellwo (2017) and measured the dynamic characteristics of the first formant (F1). They found that the speeds of F1 decreases (reflecting mouth closing movements) contained more speaker-specific information than speeds of F1 increases (reflecting mouth opening movements). Moreover, an advantage of using F1 over intensity is that F1 measures are less affected by varying distances to the microphone. This is particularly relevant in forensic scenarios; voice experts typically have no information about the mouth-to-transducer distance, and distance may vary, in an unknown way, in the course of a recording. Moreover, the result that measures of negative F1 dynamics explained more between-speaker variability than measures of positive F1 dynamics is highly congruent to He and Dellwo (2017) using intensity dynamics.

However, He et al. (2019) only focused on Zürich German read speech in laboratory settings. To evaluate the practical value of this method for forensic practices, the current research aimed to test whether the same results will be obtained using spontaneous speech, in different languages. Thus, we aimed to investigate the generalizability of the findings from He et al. (2019) to scenarios much closer to the ones found in forensic speaker comparisons.

Method

Corpora and speakers

Vocalic nuclei were manually annotated in Praat (Boersma and Weenink, 2017) in data from three corpora, in different languages. This was done using phonetic transcripts created through forced alignment of available orthographic transcripts:

- For English, telephone conversations from 14 speakers were annotated (DyVis corpus [Nolan 2011], task 2). Per speaker, between 26 and 40 sentences were included (M = 33);
- For Dutch, spontaneous face-to-face conversations from 16 gender-balanced speakers were included (Spoken Dutch Corpus <http://lands.let.ru.nl/cgn/ehome.htm>). Per speaker, between 25 and 43 sentences were included (M = 34);

– For Zürich German, the TEVOID (Dellwo et al. 2015) corpus was used, containing 16 gender-balanced speakers. Per speaker, 16 spontaneous sentences were extracted from an interview with an experimenter.

Acoustic and statistical analysis

The trajectories of F1 of each syllable nucleus were extracted using Praat (Boersma & Weenink, 2017), and the F1 dynamics (F1[+] and F1[-]) were calculated following the procedure described in He et al. (2019). The distributional characteristics of F1[+] and F1[-] in each sentence were calculated in terms of the mean (mean_F1[+] and mean_F1[-]), the standard deviation (stdev_F1[+] and stdev_F1[-]) and pairwise variability index (pvi_F1[+] and pvi_F1[-]). Multinomial logistic regressions were used to test the amount of between-speaker variability each of these measures can explain. This procedure was repeated for each of the languages.

Data processing and analysis are currently under way. We will present and discuss the results at the conference.

Acknowledgements

This work is being supported by an IAFPA research grant and an NWO VIDI grant (276-75-010).

References

- Boersma, P; Weenink, D (2017) “Praat: doing phonetics by computer,” Version 6.0.28, downloaded from <http://www.fon.hum.uva.nl/praat/>.
- Dellwo, V; Leemann, A; Kolly, M-J (2015) “Rhythmic variability between speakers: Articulatory, prosodic, and linguistic factors,” *Journal of the Acoustical Society of America* 137: 1513–1528.
- He, L; Dellwo, V (2014) “Speaker idiosyncratic variability of intensity across syllables,” in *Interspeech 2014*, Singapore, pp. 233–237.
- He, L; Dellwo, V (2016) “The role of syllable intensity in between-speaker rhythmic variability,” *International Journal of Speech, Language and the Law* 23: 243–273.
- He, L; Dellwo, V (2017) “Between-speaker variability in temporal organizations of intensity contours,” *Journal of the Acoustical Society of America* 141: EL488–EL494.
- He, L; Zhang, Y; Dellwo, V (2019) “Between-speaker variability and temporal organization of the first formant” *Journal of the Acoustical Society of America* 145: EL209–EL214.
- Leemann, A; Kolly, M-J; Dellwo, V (2014) “Speaker-individuality in suprasegmental temporal features: Implications for forensic voice comparison,” *Forensic Science International* 238: 59–67.
- McDougall, K (2006) “Dynamic features of speech and the characterisation of speakers: Towards a new approach using formant frequencies,” *International Journal of Speech, Language and the Law* 13: 89–126.
- Nolan, F (2011) *Dynamic Variability in Speech: a Forensic Phonetic Study of British English, 2006–2007* [data collection], UK Data Service.



**Universität
Zürich** UZH