29th Conference of
The International Association for Forensic Phonetics and Acoustics
22nd August – 25th August 2021, Marburg

# preliminary abstract booklet for posters

# IAFPA 2021
## 22nd August – 25th August 2021, Marburg
## – Final Program –

### Sunday , 22th August 2021

| | | | |
|---|---|---|---|
| 18:00 | Registration + Welcome | | |
| | Crash course Gathertown and Webex | | |
| 19:00 | The Tonight Show starring Gea de Jong-Lendle hosting two mystery guests | | |
| Open End | Gathertown remains open throughout the conference | | |

### Monday , 23rd August 2021

| | | | |
|---|---|---|---|
| 08:45 | Crash course Webex | | |
| **09:00** | **Welcome** | Chair | Gea de Jong-Lendle |
| | | Director Research Centre Deutscher Sprachatlas , Philipps-University Marburg | Prof Alfred Lameli |
| | | Senior Advisor for Serious Crime - Hessisches Landeskriminalamt | Kriminalhauptkommissar Klaus Lochhas |
| **09:30** | **Keynote 1** | **Michael Jessen** | **Two issues on the combination between automatic and auditory-acoustic methods in forensic voice comparison** |
| 10:30 | Talk 1 | Peter French, Jessica Wormald, Katherine Earnshaw, Philip | An Empirical Basis for System Validation and Proficiency Testing in the UK |
| **10:55** | | BREAK | |
| 11:30 | Talk 2 | Christin Kirchhuebel and Georgina Brown | Competency Testing: Opportunities and Challenges |
| 11:55 | Talk 3 | Richard Rhodes | Project proposal for IAFPA-led collaboration on method testing and validation |
| 12:20 | Talk 4 | Sula Ross, Georgina Brown and Christin Kirchhübel | Voicing Concerns: Data protection principles and forensic speech science |
| **12:45** | | LUNCH | |
| 13:45 | | **Poster session 1** | |
| **14:45** | | BREAK | |
| 15:15 | Talk 5 | Anil Alexander, Finnian Kelly and Erica Gold | A WYRED connection: x-vectors and forensic speech data |
| 15:40 | Talk 6 | Bruce Xiao Wang and Vincent Hughes | System performance as a function of score skewness, calibration and sample size in forensic voice comparison |
| 16:05 | Talk 7 | Zhenyu Wang and John H.L. Hansen | Impact of Naturalistic Field Acoustic Environments on Forensic Text-independent Speaker Verification System |
| **16:30** | | BREAK | |
| 16:50 | Talk 8 | Tomáš Nechanský, Tomáš Bořil, Alžběta Růžičková, Radek Skarnitzl and Vojtěch Skořepa | The effect of language and temporal mismatch on LTF and ASR analyses |
| 17:15 | Talk 9 | Linda Gerlach, Tom Coy, Finnian Kelly, Kirsty McDougall and Anil Alexander | How does the perceptual similarity of the relevant population to a questioned speaker affect the likelihood ratio? |
| 17:45 | Zumba (15Min) | Dr. Zumba | Something for your mental and physical wellbeing: Get those old bones moving! |

### Tuesday , 24th August 2021

| | | | |
|---|---|---|---|
| 09:30 | Talk 10 | Conor Clements, Deborah Loakes and Helen Fraser | Forensic audio in context: enhancement, suggestibility, and listener aptitude for identifying speakers in indistinct audio |
| 09:55 | Talk 11 | Valeriia Perepelytsia, Thayabaran Kathiresan, Elisa Pellegrino | Does audio recording through video-conferencing tools hinder voice recognition |
| 10:20 | Talk 12 | Camryn Terblanche, Philip Harrison and Amelia Gully | Performance of humans in detecting spoofed speech in degraded conditions |
| **10:45** | | BREAK | |
| 11:15 | Talk 13 | Luke Carroll | Bringing rhythm measures to spontaneous speech through frequently occurring speech units |
| 11:40 | Talk 14 | Kirsty McDougall, Alice Paver, Francis Nolan, Nikolas Pautz, Harriet Smith and Philip Harrison | Phonetic correlates of listeners' judgements of voice similarity within and across accents. |
| **12:05** | **Keynote 2** | **Phil Rose** | **Applications of the likelihood ratio framework in forensic speech science** |
| **12:55** | | LUNCH | |
| 14:00 | | **Poster session 2** | |
| **15:00** | | BREAK | |
| 15:30 | Talk 15 | Linda Gerlach, Kirsty McDougall, Finnian Kelly and Anil Alexander | How do Automatic Speaker Recognition systems 'perceive' voice similarity? Further exploration of the relationship between human and machine voice similarity ratings |
| 15:55 | Talk 16 | Willemijn Heeren and Lei He | Between-speaker variability in segmental F1 dynamics in spontaneous speech |
| 16:20 | | **Annual General Meeting** | |
| **18:00** | | CONFERENCE DINNER | |
| **19:30** | **Keynote 3** | **Yulia Oganian** | **Encoding and decoding of speech sounds using direct neural recordings from human auditory** |

### Wednesday , 25th August 2021

| | | | |
|---|---|---|---|
| 09:30 | Talk 17 | Helen Fraser | Updating the Likelihood Ratio debate: Behind the scenes in three Australian trials |
| 10:00 | Talk 18 | Tina Cambier-Langeveld | Speaking of authorship–can text analysis techniques be applied in forensic speaker comparison casework? |
| 10:25 | Talk 19 | Vincent van Heuven and Sandra Ferrari Disner | What's in a name? On the phonetics of trademark infringement |
| 10:50 | Talk 20 | Honglin Cao and Xiaolin Zhang | A Survey on Forensic Voice Comparison in Mainland China |
| **11:25** | | BREAK | |
| 11:40 | Talk 21 | Alice Paver, David Wright and Natalie Braber | Accent judgements for social traits and criminal behaviours: ratings and implications |
| 12:05 | Talk 22 | Kirsty McDougall, Nikolas Pautz, Harriet Smith, Katrin Müller-Johnson, Alice Paver and Francis Nolan | An investigation of the effects of voice sample duration and number of foils on voice parade performance. |
| 12:30 | Talk 23 | Paula Rinke, Mathias Scharinger, Kjartan Beier, Ramona Kaul, Tatjana Schmidt and Gea de Jong-Lendle | The effect of Angela Merkel on right temporal voice processing – an EEG study |
| **13:00** | | CONFERENCE FAREWELL | |

# Effects of speech channel on acoustic measurements and speaker discrimination from /s/

*Laura Smorenburg and Willemijn Heeren*
*Leiden University Centre for Linguistics, Leiden, The Netherlands*
b.j.l.smorenburg@hum.leidenuniv.nl, w.f.l.heeren@hum.leidenuniv.nl

Sibilant fricative /s/ is reported to be affected by various social factors such as social class (Stuart-Smith, 2007), gender and sexual orientation (e.g. Munson et al., 2006), and region (Ditewig et al., 2019). In terms of its speaker-specificity and possible potential for forensic speaker comparisons, /s/ seems to carry a useful amount of speaker information, generally ranking with nasals amongst the consonants (Kavanagh, 2012; Van den Heuvel, 1996).

Forensic casework often has to deal with scarce, or low-quality speech material such as wire-tapped telephone conversations. Although the effect of the telephone band on speech acoustics has been studied for vowels (e.g. Byrne & Foulkes, 2004), the effect on various consonants is not well documented. The effect of the telephone bandwidth (300 – 3400 Hz for landlines) is expected to be larger for some consonants than for vowels, as many consonants can have their main spectral peaks around or below the lower cut-off (nasal consonants) or above (sibilant fricatives) the upper cut-off of the telephone band. In this study, we investigate how spectral measures from /s/ are affected by the telephone band and how this affects the strength-of-evidence in speaker comparisons.

## Methodology

**Material.** Materials were taken from Task 2 in the WYRED corpus (Gold et al., 2018). In this task, speakers conversed over the telephone with an "accomplice" whilst being simultaneously recorded over a high-quality microphone (44.1 Hz sampling frequency). The landline telephone signal was wiretapped and – as is typical for these signals – only had spectral energy from 0 – 4 kHz with attenuated frequencies below 300 and above 3,400 Hz.

For 43 speakers from the Wakefield region, /s/ tokens were automatically queried from forced-aligned annotations that were generated from the orthographic transcriptions. Per speaker, the segmentations for the first one hundred /s/ tokens were then manually corrected. Tokens with overtalk, laughter, or background noise were excluded, resulting in 5,518 tokens.

**Acoustic analysis.** Following Koenig et al. (2013), a low-frequency cut-off of 550 Hz was applied across recording types to exclude ambient noise and potential lower harmonics from intruding voicing during the fricative. A high-frequency cut-off of 8 kHz was applied to the studio recordings to ensure that the highest-amplitude peak is likely to be the main front-cavity resonance (cf. Koenig et al. 2013). For the telephone recordings, an upper cut-off of 3.4 kHz was used, thus excluding attenuated frequency regions. Spectral moments (M1: mean, M2: standard deviation, M3: skewness, and M4: kurtosis) were measured automatically over the mid 50% of each /s/ token in Praat (Boersma, 2001). M1 was also measured dynamically in 20% non-overlapping windows over the full duration of the fricative. A cubic polynomial fit resulted in four coefficients: an intercept and three dynamic coefficients. The cubic fit was a better fit to the data than the quadratic fit: $\chi^2(1) = 27.79$, $p < .001$.

**Statistical analysis.** The statistical analysis had two parts. First, linear mixed-effects

modelling was used to examine the fixed effect of Channel (sum-coded). Second, LR analysis was used to calculate the strength-of-evidence in speaker comparisons. Highly correlating features (r > .50) were excluded (M3 and the dynamic intercept of M1 were excluded for their high correlation with M1). Given that only 43 speakers had been annotated so far, a leave-one-out MVKD implementation with calibration was used (Morrison, 2007).

## Results

As expected, large differences in acoustic measurements were found by speech channel, with much lower values in the telephone band (see e.g. M1 and M2 in Table 1). Only the linear and cubic terms from the M1 dynamics did not show significant channel effects, indicating only minimal effects of channel on the dynamics.

**Table 1.** Best-fitting mixed-models for M1 and M2

| | | | M1 | | | | M2 | | |
|---|---|---|---|---|---|---|---|---|---|
| Random effects | | | Var | SD | Cor | | Var | SD | Cor |
| Speaker | (intercept) | | 99,246 | 315 | | | 8,833 | 94 | |
| | Channel | | 110,887 | 333 | -0.46 | | 8,975 | 95 | -0.71 |
| Residual | | | 250,756 | 501 | | | 39,192 | 198 | |
| Fixed effects | | Est. | SE | t | *p* | Est. | SE | t | *p* |
| (intercept) | | 3,363 | 48 | 69.6 | <.001 | 981 | 14 | 67.8 | <.001 |
| Channel: telephone | | –1,622 | 51 | –31.8 | <.001 | –298 | 15 | –20.4 | <.001 |

Although there were effects of channel on the acoustics, when suspect and offender data match on speech channel, its effect on the strength-of-evidence in speaker comparisons seems to be minimal. This suggests that the lower frequencies might capture similar speaker information as the higher frequencies.

**Table 1.** Same-speaker and different-speaker LLRs and system performance measures

| | LLR SS | LLR DS | Cllr | Cllr-min | EER% |
|---|---|---|---|---|---|
| Studio | 3.08 | -11.40 | 0.27 | 0.22 | 6.2 |
| Telephone | 3.23 | -16.88 | 0.24 | 0.20 | 6.3 |

## References


Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glot Int.*, *5*(9/10), 341–347.

Byrne, C., & Foulkes, P. (2004). The "Mobile Phone Effect" on Vowel Formants. *Speech Lang. and the Law*, *11*(1), 83–102.

Ditewig, S., Pinget, A. C. H., & Heeren, W. F. L. (2019). Regional variation in the pronunciation of /s/ in the Dutch language area. *Nederlandse Taalkunde*, *24*(2), 195–212.

Gold, E., Ross, S., & Earnshaw, K. (2018). The "West Yorkshire Regional English Database". *Proc. Interspeech*, *2018-September*, 2748–2752.

Kavanagh, C. M. (2012). *New consonantal acoustic parameters for forensic speaker comparison* University of York.

Morrison, G.S. (2007). Matlab implementation of Aitken & Lucy's (2004) forensic likelihood-ratio software using multivariate-kernel-density estimation.

Munson, B., McDonald, E. C., DeBoe, N. L., & White, A. R. (2006). The acoustic and perceptual bases of judgments of women and men's sexual orientation from read speech. *J.Phon.*, *34*, 202–240.

Stuart-Smith, J. (2007). Empirical evidence for gendered speech production: /s/ in Glaswegian. *Change in Phon.: Papers in Lab. Phon. 9*, 65–86.

Van den Heuvel, H. (1996). *Speaker variability in acoustic properties of Dutch phoneme realisations* Radboud Universiteit.