Looking into the real world: 
Likelihood ratio variability under forensically realistic conditions

Yuko Kinoshita

School of Culture, History and Language, the Australian National University, Australia

Yuko.Kinoshita@anu.edu.au

Abstract

This study set out to investigate how the speech of a single speaker can vary depending on their interlocutor and apparent emotional status and, consequently, how this affects likelihood ratio (LR)-based forensic voice comparison, using authentic data from past forensic casework. The results revealed that these factors do have a significant impact on the LR calculations, and voice comparisons between the testing data with mismatched conditions produce much less reliable results than those recorded under matching conditions.

Index Terms: forensic voice comparison, emotion, interlocutor, within-speaker variability, likelihood ratio

1. Introduction

Forensic voice comparison is a complex and challenging task. The research in this field has made significant progress over the last decade. An LR-based framework has been implemented, both in phonetics-based approaches and in engineering ones (e.g. [1-3]). Various features have been tested (e.g. [4-7]), LR calculation methods have been improved (e.g. [8-10]), and different approaches to evaluate the quality of LR and to post process have been put forward (e.g. [11-14]). For a comprehensive overview of these and other developments, see [15]). When it comes to casework situations, however, forensic scientists still face many problems, which are yet to be fully studied.

One of the difficulties with forensic voice comparison is the lack of control over testing data. In casework situations, scientists often have to compare two (or more) voice recordings made under very different circumstances. Recordings of unknown speakers (usually from a crime scene) are likely to be recorded in a noisy environment, with the speaker emotionally aroused—perhaps even shouting—and talking in a casual style to someone they know. Recordings of known speakers are quite often made at interviews between police officers and an apprehended suspect. The mood of these speakers tend to be nervous or depressed or at least unhappy; and they often use a different speaking style from when they talk normally to people they know. We expect these different conditions to influence speech acoustics, but we still don’t know to what extent this affects the LR-based forensic voice comparison. “Are those recordings really comparable?” is almost the first question forensic scientists should ask themselves when faced by casework data, but it is rare that we can answer with full confidence. Should we just refuse to do any analysis, if we are not sure about the comparability; or is there still something that we can usefully do? Where is the useful limit for the recording conditions to be regarded as comparable? These questions must be answered, or at least be considered carefully.

Being motivated by the author’s experiences in voice comparison casework as a forensic scientist, this study revisits a set of past casework data with permission from the relevant bodies. The dataset consists of 15 telephone calls made by a known speaker, talking to various interlocutors in varying emotional states. Although what we can conclude from studying a single speaker is limited, this study aims to provide a useful starting point for tackling the questions presented above.

2. Procedure

2.1. Data

2.1.1. Recordings

The testing data for this study consists of 15 phone calls made by an adult male of known identity. These calls were made from an Australian remand centre (detention centre for unconvicted suspects) over three days. They are completely spontaneous, and the physical circumstances of the speaker were consistent across these 15 phone calls. Other conditions such as the speaker’s speaking style, emotional state and interlocutor varied widely.

The duration of the phone calls ranged from 53 to 720 seconds, altogether over 100 minutes. A substantial proportion was the target speaker’s speech. The target speaker was bilingual in Australian English and Arabic, and switched freely between the two languages. This pilot study limits the analysis to the sections spoken in English. His English did not have discernable characteristics typical of a non-native speaker of English.

The recordings were labelled for interlocutors, and emotional states, as well as for word boundaries and segments. Interlocutors influence our speaking style (e.g. [16]), and hence speech acoustics. The target speaker talked with a range of interlocutors including family members and other male and female speakers who seem to have a close personal connection with him, such as friends, cousins and co-workers.

Emotional states were labelled based on the author’s auditory impressions of the speaker’s perceived emotional states. Using auditory impressions from a panel of listeners is a scientifically rigorous approach, but the agreements for the data use did not allow this. The author’s auditory impressions were supplemented by additional clues from linguistic and non-linguistic cues, such as laughter, yelling, and appearance of strongly abusive words. The labels used in classifying emotional states and interlocutors are summarized in Table 1.

This categorization of the emotional states may appear too fine grained. They do not reflect author’s confidence in such classification, but they are to avoid forced categorization. This way, groups whose classification appears unreliable can be excluded from the analysis at a later stage.
Table 1. Summary of the variables

<table>
<thead>
<tr>
<th>Variables</th>
<th>Types</th>
</tr>
</thead>
<tbody>
<tr>
<td>Emotional states</td>
<td>happy, friendly, neutral/friendly, neutral, neutral/irritated, irritated, irritated/angry (in the order of positiveness of the emotion)</td>
</tr>
<tr>
<td>Interlocutors</td>
<td>mother, father, wife, child, male, female</td>
</tr>
</tbody>
</table>

2.1.2. Feature extraction

The acoustic quality of the recordings was sufficient for formant extraction for most part, but telephone bandwidth effects were observed. The power spectrum revealed that spectral energy was suppressed in the region below 300Hz and above 3400Hz. Therefore F1 of some high vowels would be affected.

F1–F3 of 13 phonemes of Australian monophthongs (/æ/, /ɛ/, /ɛː/, /ɪ/, /ɒ/, /ɔ/, /ɒː/, /ʊ/) were extracted. Praat was used for tagging of the target vowels and extracting formants. F1 to F3 were sampled at the mid point of the vowel duration semi-automatically, and the values were checked against expected values for those formants, referring to their distributions in preceding studies [17-19]. Where the extracted values significantly deviated from the expected range, they were re-measured and manually corrected as required. The vowel space of this speaker is presented in Figure 1. Each vowel indicates the mean, and the ellipses mark the 95% range of the distribution. Each vowel spreads widely compared to data such as Bernard’s [17] in which the utterances were much more controlled. F2 of /ɔː/ varied particularly widely and it had a bi-modal distribution.

![Figure 1 Target speakers vowel plot](image)

2.2. Statistical analysis

The dataset was firstly analysed by multivariate analysis of variance to examine whether the two factors, emotional state and interlocutors, have statistically significant effects. The formant/vowel combinations found to be affected in question were further analysed.

Then multivariate kernel density likelihood ratios (MVKD) [9] were calculated for the testing samples with various mismatched conditions, and the effects of the different interlocutors and emotional states were examined. Due to the lack of data comparable to that in this study, no calibration was performed.

For the background population, this study used Bernard’s formant data [17]. This dataset was extracted from recordings made under strictly controlled studio conditions in the 1960s, clearly not ideal for evaluating the strength of evidence for the testing data used here. In general, mismatch of conditions between a background population and test data makes the testing samples seem more atypical, and is likely to produce more extreme values as the significance of their similarity (or dissimilarity) will be overrated. For a same-speaker comparison, we expect a log$_{10}$LR (LLR) greater than 0. Since this study aims to examine under which conditions the LRs deviate from this expectation rather than examining the strength of evidence itself, Bernard’s formant dataset was used for its size and availability.

For the calculation of LRs, the dataset was divided into groups based on two factors: interlocutors and emotional states. For the examination of the effect of emotional states, five subgroups were selected: neutral/friendly, neutral, neutral/irritated, irritated, and irritated/angry, based on the availability of sufficient vowel data (183, 262, 257, 190, and 331 respectively). They were compared against each other, resulting in 10 different emotional state combinations.

For interlocutor effects, we selected three speakers: wife, male friends (which includes multiple individuals), and his young child. Again, these were selected because they had sufficient numbers of vowels (757, 454, and 31 respectively).

To have baseline LRs, the datasets with matching conditions were also compared. Interlocutor wife was selected for this purpose, as it had the largest amount of vowel data. This dataset consisted of four different emotional states: neutral, neutral/irritated, irritated, and irritated/angry; from neutral to more negative and strong. These emotion groups were further divided into two, and these two sets were compared to observe the LRs where the testing data’s conditions were strictly matched.

In all cases, LRs were calculated for each vowel separately, using measurements for F1, F2 and F3; except for /iː/ and /ʊː/, where F1 was excluded from analysis as telephone bandwidth would have affected it.

3. Results

3.1. ANOVA results

Of 13 vowels, /ɒ/ was excluded from the analyses as there were only 3 tokens. Table 2 summarises the results of ANOVA, revealing that the factors in question affected 11 out of 12 vowels. No statistically significant effect on /ɛː/ was found. Thus the subsequent analysis will focus on 11 vowels excluding /ɔː/ and /ɛː/.

![Table 2. Summary of ANOVA](image)
3.2. LR testing

3.2.1. Initial observations

Figure 2 below presents the 90% distributions range of the obtained log_{10}LRs (LLRs). The results for all vowels were combined. The black line shows emotional state mismatch; the red line shows interlocutor mismatch; and the blue lines show matching conditions. The x-axis has been truncated for ease of observation.

Perhaps due to the significant difference in the recording conditions between the test data and Bernard’s data, some alarmingly large LLRs were observed. They supported the hypothesis consistent with the identity of the speaker (the same-speaker hypothesis), but their size suggests they are unreliable. Where a suitable training dataset for calibration is available, such spurious values may be less of a problem. However, the fact that the comparisons with matching conditions produced no such values may be noteworthy. Mismatch of conditions in both emotional states and interlocutor appears to introduce unpredictability in the LR calculation.

Also, although the three distributions substantially overlap with one another, the locations of their peaks suggest that comparisons under matching conditions tend to produce on balance slightly better evidence than the ones with condition mismatch.

![Figure 2: Distribution of LLRs for the three different comparison conditions (emotion mismatch, interlocutor mismatch, matching)](image)

3.2.2. Effect of Emotional status

Figure 3 presents the ranges of LLR produced by various combinations of emotional states. For comparison it also shows LLRs for matching combinations. LLRs were calculated for 10 mismatch combinations and four matched combinations.

The results revealed no systematic relationship between LLRs and the types, intensity or degree of mismatch in emotional state. For instance, Combination 5 in Figure 3 (irritated vs neutral_irritated) produced the strongest LLRs with a strikingly wide variation. On the other hand, other combinations that produced larger LLRs (Combinations 4, 6, 9 and 10) varied in quality and intensity of emotion, and no commonality was found. Also Combination 2 (irritated_angry vs neutral_friendly), which was in the greatest emotional mismatch did not produce notable results. The only tendency found was that the comparisons between the samples whose emotional states were not matched tend to produce a wider range of LLRs (possibly spurious ones,) compared those with matched emotional states such as comparisons 11–14.

![Figure 3: Comparisons of the range of LLRs based on emotional states](image)

3.2.3. Effect of Interlocutors

The results for the interlocutor mismatch comparisons are presented in Figure 4. These did not produce as wide a range of LLRs as the emotion mismatch. Within this limited variability, however, the comparisons involving child as the interlocutor produced wider ranges of LLRs. There are two possible explanations for this: speaking styles and data size. The target speaker talked with his child in a clearly different speaking style. Also there were much longer recordings for the conversation between him and his wife or male friends than him and his child. So the amount of available data was very different. This effect has to be re-examined with equally sized datasets.

![Figure 4: Comparisons of the range of LLRs based on Emotional status](image)
4. Summary and discussion

Using authentic data from forensic casework as the testing data, this study conducted a pilot investigation on how LR-based forensic voice comparison is affected when two testing samples have a mismatch of emotional states and interlocutors. The results revealed that forensic voice comparisons with such mismatches tend to produce wider variations of LLRs. Although no systematic relationship emerged in terms of the relationship between LLRs and the types of mismatch, this finding proves the need for further study into the mechanisms at work and which effects are most significant, as well as cautioning forensic scientists working in voice comparison about the unpredictable effects of mismatched conditions.

As for future tasks, firstly the formant data for individual vowels need to be re-examined for each specific condition. As discussed above, the classification of the emotional states was far from ideal in this study. A statistically based method needs to be sought for this task. Also, comparisons with other speakers of the same variety of Australian English will be useful for separating potentially influential factors, such as ethnocultural, individual characteristics of speakers, emotional states and interlocutor.

Re-analysing against a different background population such as [20] or [21] would be also useful. While they are far from having matching conditions to the testing data used here, at least they are contemporary and spontaneous. Bernard’s dataset is useful for pilot experiments and has been used in many previous studies, but the results obtained here seem to suggest that the differences in the data characteristics may have been too great. This study observed a significant number of clearly spurious LRs. These appear to be caused by the condition mismatch between the two testing recordings, but also possibly aggravated by the choice of background data. While the absolute values of LLRs were not the focus of this study, these outliers obscured the overall picture.

The number of samples used in the LR calculation needs to be controlled more. In the first instance, this study used all available data for the analysis, but the differing numbers of data points used for the LR calculation might have obscured the relationship that was the focus of this study.

Further, this study limited the analyses to formant data. Other features such as F0 and temporal features could provide other important cues. This pilot study could not gain enough understanding to develop techniques to overcome the mismatched conditions in casework. However, it at least demonstrated that condition mismatch between testing samples can have a serious impact on LR-based voice comparison. Continuing studies are essential if we are to aspire to contribute to real life casework.

5. Acknowledgements

I would like to thank my three anonymous reviewers for their constructive comments and feedback. However, for any remaining shortcomings found in this paper, I am solely responsible.

6. References


Forensics 1

10:25-12:05, 3rd December 2014, Session Chair: Viktoria Papp

16 LR-based forensic comparison under severe test-data scarcity.
Yuko Kinoshita\textsuperscript{1}, Michael Wagner\textsuperscript{1}

\textsuperscript{1}University of Canberra

Replicate mismatch between test and background/development databases: The effect on the performance of likelihood ratio-based forensic voice comparison
Shunichi Ishihara\textsuperscript{1}

\textsuperscript{1}Australian National University

Regional variation and the definition of the relevant population in likelihood ratio-based forensic voice comparison using cepstral coefficients
Vincent Hughes\textsuperscript{1}, Paul Foulkes\textsuperscript{1}

\textsuperscript{1}University of York

Comparison between speech parameters for forensic voice comparison using mobile phone speech
Esam Alzqholi\textsuperscript{1}, Balamurali B T Nair\textsuperscript{1}, Bernard Guillemin\textsuperscript{1}

\textsuperscript{1}University of Auckland
Perception 2

14:30-15:20, 4th December 2014, Session Chair: Lynn Clark

Improvements to vowel categorization in non-native regional accents resulting from multiple- versus single-talker training: A computational approach

Sarah Wright¹, Jason Shaw¹, Catherine Best¹, Gerard Docherty², Bronwen Evans³, Paul Foulkes⁴, Jen Hay⁵, Karen Mulak¹

¹University of Western Sydney, ²Griffith University, ³University College London, ⁴University of York, ⁵University of Canterbury

Thai phonetically balanced word recognition test: Test-retest reliability and error analysis

A. Munthuli¹, C. Tantibundhit¹, C. Onsuwan¹, K. Kosawat²

¹Thammasat University, ²Thailand National Electronics and Computer Technology Center

Forensics 2

15:45-16:10, 4th December 2014, Session Chair: Paul Foulkes

A first attempt at compensating for effects due to recording-condition mismatch in formant-trajectory-based forensic voice comparison

Ewald Enzinger¹

¹University of New South Wales

Looking into the real world: LR variability under forensically realistic conditions

Yuko Kinoshita¹

¹Australian National University

Speech Production 1

14:30-15:20, 4th December 2014, Session Chair: Catherine Theys

Temporal planning in the production of Australian English compounds

Ivan Yuen¹, Nan Xu Rattanasone¹, Gretel McDonald¹, Rebecca Holt¹,

¹Macquarie University

Time spent talking in retirement

Nina Fhärm¹, Frida Wigell Skoglund¹, Jan van Doorn¹

¹Umeå University

New Technologies 2

10:10-11:50 5th December 2014, Session Chair: Catherine Watson

Automatic detection of speech truncation and speech rate

Chung Ting Justine Hui¹, Teh June Chin¹, Catherine Watson¹

¹University of Auckland
Short utterance PLDA speaker verification using SN-WLDA and variance modelling techniques

Ahilan Kanagasundaram1, David Dean1, Sridha Sridharan1
1Queensland University of Technology

Rescaling clustering trees using impact ratios for robust hierarchical speaker clustering

Houman Ghaemmaghami2, David Dean1, Shahram Kalantari1, Sridha Sridharan1
1Queensland University of Technology

Phonetic spoken term search using topic information

Shahram Kalantari1, David Dean1, Sridha Sridharan1
1Queensland University of Technology

Speech Production 2

10:10-11:50 5th December 2014, Session Chair: Beth Hume

Retention of Spanish coda /s/ by speakers of Kashibo-Kakataibo

Sally Bowman1, Roberto Zariquiey Biondi2, Marija Tabain3
1La Trobe University, 2Pontificia Universidad Catolica del Peru

Assibilation in Trans-New Guinea languages of the Bird's Head region

Fanny Cottet1
1Australian National University

Iconicity in Korean consonantal symbolism

Nahyun Kwon1
1University of Queensland

Some initial findings regarding first language influence on playing brass instruments

Matthias Heyne1, Donald Derrick1,2
1University of Canterbury, 2University of Western Sydney

Prosody

12:40-13:55, 5th December 2014, Session Chair: James Gruber

Sociophonic and prosodic influences on judgements of sentence type

Paul Warren1
1Victoria University of Wellington

Stress-meter alignment in American hip hop

Casey Tait1, Marija Tabain1, Ingrid Sykes1
1La Trobe University

Examining the influence of pitch accents on word learning in German

Michael Walsh1, Katrin Schweitzer1, Hinrich Schütze2, Dermot Lynott3
Perception 3
12:40-13:55, 5th December 2014, Session Chair: Michael Tyler

A restriction on minimal words, or word endings?
197
Jason Brown1, Forrest Panther1
1University of Auckland

Orthographic effects on phonetic cue weighting
201
Grant McGuire1
1University of California Santa Cruz

Frequency in the input affects perception of phonological contrasts for native speakers
205
Rikke Louise Bundgaard-Nielsen1, 2, Brett Baker3
1La Trobe University, 2University of Western Sydney, 3University of Melbourne

Perception 4
14:20-15:35 5th December 2014, Session Chair: Jen Hay

Influence of phonological, morphological, and prosodic factors on phoneme detection by native and second-language adults
210
Valeria Peretokina1, Michael Tyler1, Catherine Best1
1University of Western Sydney

Perception of Italian and Japanese consonant length by native speakers of Australian English and Italian: A pilot study
214
Kimiko Tsukada1, Felicity Cox1, John Hajek2, Yukari Hirata3,
1 Macquarie University, 2University of Melbourne, 3Colgate University

Is more always better? The perception of Dutch vowels by English versus Spanish listeners
218
Samra Alispahic1, Paola Escudero1, Karen Mulak1
1University of Western Sydney

Child Language Acquisition
14:20-15:10 5th December 2014, Session Chair: Jeanette King

Consonant inventory of infants aged 0-6 months
223
Adele Gregory1, Marija Tabain1, Michael Robb2
1La Trobe University, 2University of Canterbury

Investigating the effect of intrusive noise levels on speech perception in an open-plan kindergarten classroom
227
Kiri Mealings1, Katherine Demuth1, Jorg Buchholz1,2, Harvey Dillon2
1 Macquarie University, 2National Acoustics Laboratories
Posters

The voiced emphatic coronal stop [d] in the Hadrami Arabic: A socio-phonetic study
Hamad Altari¹
¹University of Auckland

Contour-based analysis of EGG data from words in isolation and connected speech
Stephen Bier¹, Catherine Watson¹, Clare McCann¹
¹University of Auckland

Venezuelan Spanish intransitives: More prosodic than you'd think
Sasha Calhoun¹, Erwin La Cruz¹, Ana Olssen¹
¹Victoria University of Wellington

Recency effects on word-medial /t/ in New Zealand English: initial observations
Lynn Clark¹, Liam Walsh¹
¹University of Canterbury

F0 as a word boundary cue for segmenting New Zealand English ethnolects
Kylie Fitzgerald¹, Viktoria Papp¹, Jen Hay¹
¹University of Canterbury

Distinguishing dysarthric speech: Vowel acoustics and measurements
Annalise Fletcher¹ Megan McAuliffe¹, Kaitlin Lansford², Julie Liss³
¹University of Canterbury, ²Florida State University, ³Arizona State University

Ability to identify unfamiliar speech sounds negatively correlates with second language proficiency
Yurika Hashimoto¹, Ian Wilson¹, Younghyon Heo¹
¹University of Aizu

Why Indo-Aryan languages adapt English alveolars as retroflexes: Acoustic evidence from Punjabi
Qandeel Hussain¹
¹Macquarie University

I Bag Your Pardon: The Albertan ae/e shift and community grammars
Jacqueline Jones¹, Stephen Winters¹
¹University of Calgary

Diphthong trajectories in Māori
Jeanette King¹, Catherine Watson², Margaret Maclagan², Peter Keegan², Ray Harlow²
¹University of Canterbury, ²University of Auckland, ³University of Waikato (Retired)
Welcome to Delegates

Nau mai, haere mai, tauti mai ki Ōtautahi i Te Waipounamu mō tēnei hui e pā ana ki ngā āhuatanga o te reo. Tēnā koutou katoa.

A warm welcome to the 15th Australasian International Conference on Speech Science and Technology (SST 2014), in Christchurch, New Zealand.

SST held its first meeting in Canberra in 1986. Since then, the conference has remained the most important regular conference on speech science and technology held in the Southern Hemisphere. It has always been an outstanding forum for debate, discussion, sharing of results and ideas, nurturing young scholars, and networking. The New Zealand Institute of Language, Brain and Behaviour is very proud to uphold this fine tradition by hosting the event in Christchurch this year. I would like to thank all of the authors who have contributed to a simply outstanding programme of talks and posters.

I would also like to gratefully acknowledge the support of our sponsors: Tobii, Neurospec, NDI and GE Healthcare Ltd; and to thank the ASSTA committee, the many reviewers and the SST 2014 Organising Committee for their significant contributions to the organisation of this event. Enormous thanks are also due to our unparalleled conference organiser and coordinator - NZILBB Manager, Emma Parnell.

I trust that you will all enjoy the conference, and hope that you will also find some time to relax, and to explore Christchurch and its surroundings.

Conference Chair

Jen Hay
University of Canterbury

ASSTA Corporate Members

- Appen PTY Ltd
- HEARing CRC
- Cochlear PTY Ltd
- Spectral Dynamics
# SST 2014 Conference Committee Members

## Organising Committee

<table>
<thead>
<tr>
<th>Lynn Clark</th>
<th>Donald Derrick</th>
<th>Jen Hay</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beth Hume</td>
<td>Jeanette King</td>
<td>Megan McAuliffe</td>
</tr>
<tr>
<td>Emma Parnell</td>
<td>Catherine Theys</td>
<td>Paul Warren</td>
</tr>
<tr>
<td>Catherine Watson</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

## Review Board

<table>
<thead>
<tr>
<th>Kirrie Ballard</th>
<th>Jason Brown</th>
<th>Denis Burnham</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lynn Clark</td>
<td>Linda Cupples</td>
<td>Julien Epps</td>
</tr>
<tr>
<td>Paul Foulkes</td>
<td>David Grayden</td>
<td>James Gruber</td>
</tr>
<tr>
<td>Bernard Guillemin</td>
<td>John Hajek</td>
<td>Jen Hay</td>
</tr>
<tr>
<td>Colleen Holt</td>
<td>Beth Hume</td>
<td>Shunichi Ishihara</td>
</tr>
<tr>
<td>Caroline Jones</td>
<td>Jeesum Kim</td>
<td>Jeanette King</td>
</tr>
<tr>
<td>Trent Lewis</td>
<td>Deborah Loakes</td>
<td>Robert Mannell</td>
</tr>
<tr>
<td>Megan McAuliffe</td>
<td>Kirsty Dougall</td>
<td>Kevin McGowan</td>
</tr>
<tr>
<td>Sharynne McLeod</td>
<td>Peggy Mok</td>
<td>Chutamanee Onsuwan</td>
</tr>
<tr>
<td>Carmel O'Shanennessey</td>
<td>Pascal Perrier</td>
<td>David Powers</td>
</tr>
<tr>
<td>Peter Racz</td>
<td>Korin Richmond</td>
<td>Philip Rose</td>
</tr>
<tr>
<td>Mridula Sharma</td>
<td>Jason Shaw</td>
<td>Chilin Shih</td>
</tr>
<tr>
<td>Donal Sinex</td>
<td>Marija Tabain</td>
<td>Roberto Tongneri</td>
</tr>
<tr>
<td>Michael Tyler</td>
<td>Michael Wagner</td>
<td>Paul Warren</td>
</tr>
<tr>
<td>Catherine Watson</td>
<td>Kevin Watson</td>
<td>Maria Klara Wolters</td>
</tr>
</tbody>
</table>

## ASSTA Executives

<table>
<thead>
<tr>
<th>Denis Burnham</th>
<th>Shunichi Ishihara</th>
<th>Marija Tabain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Roberto Togneri</td>
<td>Catherine Watson</td>
<td></td>
</tr>
</tbody>
</table>
Invited Workshop Speakers

Donald Derrick (University of Canterbury & University of Western Sydney)
Ultrasound for speech
2nd December 2014, 9:30-12:30

Donald Derrick is a joint Post-Doctoral Fellow at the New Zealand Institute of Language Brain and Behaviour, University of Canterbury and the MARCS Institute at the University of Western Sydney. He studies speech production and perception in order to identify the phonetic constraints on low-level speech production, and the low-level percepts that can enhance or interfere with speech perception. Donald's research has implications for speech planning, constraints on speech production, and multi-modal speech perception, including the possibility of enhancing speech perception among the hard-of-hearing.

T. Florian Jaeger (University of Rochester)
Regression and mixed effects models
2nd December 2014, 14:00-17:00

T. Florian Jaeger is an Assistant Professor at the Department of Brain and Cognitive Sciences and the Department of Computer Science at the University of Rochester. He is interested in to what extent the human language production and comprehension systems are organised to be efficient. He uses psycholinguistic experimentation and statistical modeling to investigate speakers' choice at various linguistic levels of online language production and the pressures driving these decisions. His research also asks what cues and what methods language users employ when estimating probabilities of words and structures.

Chigusa Kurumada (University of Rochester) and Meredith Brown (Tufts University)
Perception and interpretation of speech prosody
2nd December 2014, 9:30-12:30

Chigusa Kurumada is an Assistant Professor at the Department of Brain and Cognitive Sciences at the University of Rochester. She is interested in the way adults and children find coherent patterns in input signals and induce linguistic representations. Her recent work focuses on roles of the adaptive nature of the human cognitive system in language comprehension, investigating how it navigates variability in prosodic information and achieves robust pragmatic inferences. Bridging topics on child language acquisition and psycholinguistics, she explores the mechanism of language learning and use throughout the lifespan.

Invited Plenary Speakers

Gerry Docherty (Griffith University)
Stitching together the story on sociophonetic variation: Advances and challenges
3rd December 2014, 9:00-10:00

Gerry Docherty is Professor and Dean (Research) in the Arts, Education & Law academic group at Griffith University. A common strand through all Gerry’s research work has been a focus on quantitative acoustic analysis of aspects of speech with a view to enhancing understanding of the nature of phonetic variability and its implications for phonetic theory. While much of his work has been focused on normal adult speakers, he has also investigated the acquisition of speech sound patterning in children and the nature of speech in populations of speakers with impaired speech
production. He is currently Chief Investigator on the ARC-funded Discovery Project entitled *The social dynamics of language: a study of phonological variation and change in West Australian English* (2013-2015). He is also currently working on a second ARC-funded Discovery Project entitled *You came TO DIE?! Perceptual adaptation to regional accents as a new lens on the puzzle of spoken word recognition* (2012-15).

**Gail Gillon (University of Canterbury)**

*Phonological awareness development in children with speech sound disorder*

4th December 2014, 13:30-14:30

Gail Gillon is a Professor at and Pro-Vice Chancellor of the College of Education at the University of Canterbury. She is also a member of the Ngāi Tahu Research Centre, the New Zealand Institute of Language, Brain and Behaviour and Co Director of the Language and Literacy Research Lab. Gail’s research focuses on understanding the relationship between developmental spoken and written language disorders. Her active research includes: the prevention of reading disorder in children with speech and language impairment; phonological awareness intervention; the development of phonological representations in children with unintelligible speech; conversational and narrative language development in New Zealand children; childhood apraxia of speech; phonological awareness development in children with Down syndrome; speech and language development in young children with traumatic brain injury; and communication impairment in children who are bilingual.

**T. Florian Jaeger (University of Rochester)**

*Speech perception and adaptation an inference under uncertainty*

5th December 2014, 8:45-9:45

See Invited Workshop Speakers

**Bastiaan Kleijn (Victoria University of Wellington)**

*A simple model of speech communication and its application to intelligibility enhancement*

3rd December 2014, 13:00-14:00

Bastiaan Kleijn is a Professor at the School of Engineering and Computer Science at Victoria University of Wellington. Bastiaan is also a part-time Professor at Delft University of Technology. His research interest is to develop solutions based on signal processing, information theory, and machine learning for problems involving audio, video, and biomedicine. Examples of topics that he has worked on recently are: distributed processing, blind source separation, spatial audio, entropy estimation, robust source coding for packet networks, single and multi-channel speech enhancement, video enhancement, camera control for multiview video, auditory modeling, audio and video coding, model estimation, and algorithms that predict perceived signal quality.
Jane Stuart-Smith (University of Glasgow)

Twa son, some soldiers, and a city: An acoustic phonetic investigation of real-time change in Scottish English

4th December 2014, 8:45-9:45

Jane Stuart-Smith is a Professor in English Language and Director of the Glasgow University Laboratory of Phonetics. She is the author of a number of articles on variation and change in Glaswegian accent, the impact of the broadcast media on language variation and change, the sociophonetics of British Asian accents, the phonetics and phonology of British varieties of Panjabi, the acquisition of literacy in Panjabi/English bilingual children, historical linguistics and sound-change. Jane is currently running two projects, one looking at how Glaswegian accent has changed since the 1970s, and the other to develop an online resource for teaching articulatory phonetics. She is also working with Jim Scobbie and Eleanor Lawson on a project to investigate further the transmission of articulatory variation, using Ultrasound Tongue Imaging to look particularly at derhoticization in Scottish English. She is also working to help Microphonics develop their innovative computer-literacy tool to help learners of English as a Second Language.
<table>
<thead>
<tr>
<th>Time</th>
<th>SESSION A-Savoy West</th>
<th>SESSION B-Savoy 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.45am</td>
<td>Welcome</td>
<td></td>
</tr>
<tr>
<td>9:00-10:00</td>
<td>Plenary: Gerry Docherty (Griffith University)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Stitching together the story on sociophonetic variation:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Advances and challenges</td>
<td></td>
</tr>
<tr>
<td>10:00-10:25</td>
<td>Morning Tea</td>
<td></td>
</tr>
<tr>
<td>10:25-10:50</td>
<td>LR-based forensic voice comparison under severe test-</td>
<td>Phonologisation of vowel duration and nasalised</td>
</tr>
<tr>
<td></td>
<td>data scarcity.</td>
<td>/æ/ in Australian English.</td>
</tr>
<tr>
<td></td>
<td>Yuko Kinoshita &amp; Michael Wagner</td>
<td>Felicity Cox &amp; Sallyanne Palethorpe</td>
</tr>
<tr>
<td>10:50-11:15</td>
<td>Replicate mismatch between test and background/development</td>
<td>Why the SQUARE vowel is the most variable in</td>
</tr>
<tr>
<td></td>
<td>databases: The effect on the performance of likelihood</td>
<td>Sydney.</td>
</tr>
<tr>
<td></td>
<td>ratio-based forensic voice comparison.</td>
<td>Nhunh Nguyen &amp; Jason Shaw</td>
</tr>
<tr>
<td>11:15-11:40</td>
<td>Regional variation and the definition of the relevant</td>
<td>Identifying /eI/-/æI/: A comparison between two</td>
</tr>
<tr>
<td></td>
<td>population in likelihood ratio-based forensic voice</td>
<td>regional Australian towns.</td>
</tr>
<tr>
<td></td>
<td>comparison using cepstral coefficients.</td>
<td>Deborah Loakes, John Hajek, Joshua Clothier &amp;</td>
</tr>
<tr>
<td></td>
<td>Vincent Hughes &amp; Paul Foulkes</td>
<td>Janet Fletcher</td>
</tr>
<tr>
<td>11:40-12:05</td>
<td>Comparison between speech parameters for forensic</td>
<td>Comparing acoustic analyses of Australian English</td>
</tr>
<tr>
<td></td>
<td>Esam Alzqhol, Balamurali B T Nair &amp; Bernard Guillemin</td>
<td>Jaydene Elvin &amp; Paola Escudero.</td>
</tr>
<tr>
<td>12:05-13:00</td>
<td>Lunch</td>
<td></td>
</tr>
<tr>
<td>13:00-14:00</td>
<td>Plenary: Bastiaan Kleijn (Victoria University of</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Wellington)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A simple model of speech communication and its application</td>
<td></td>
</tr>
<tr>
<td></td>
<td>to intelligibility enhancement</td>
<td></td>
</tr>
<tr>
<td>14:00-14:25</td>
<td>Segmental and Tonal Errors in L2 Mandarin Speech</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Produced by Australian English Learners.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Wentao Gu, Ting Zhang &amp; Chiharu Tsurutani</td>
<td></td>
</tr>
<tr>
<td>14:25-14:50</td>
<td>Does immersion experience reduce /r/-/J/ category</td>
<td></td>
</tr>
<tr>
<td></td>
<td>overlap for Japanese learners of English?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Michel Tyler, Saya Kawase &amp; Mark Antoniou</td>
<td></td>
</tr>
<tr>
<td>14:50-15:15</td>
<td>Afternoon Tea</td>
<td></td>
</tr>
<tr>
<td>15:15-15:40</td>
<td>Tonal alignment of focal pitch accents in two varieties</td>
<td></td>
</tr>
<tr>
<td></td>
<td>of Indian English.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Olga Maxwell &amp; Janet Fletcher</td>
<td></td>
</tr>
<tr>
<td>15:40-16:05</td>
<td>Amplitude and F0 as acoustic correlates of Kelantan</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Malay word-initial geminates.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mohd Hilmi Hamzah, Janet Fletcher &amp; John Hajek</td>
<td></td>
</tr>
<tr>
<td>16:05-16:30</td>
<td>Pitch accents and prosodic properties of the clitic in</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Yukulta (Tangkic).</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cicely Bonnin</td>
<td></td>
</tr>
<tr>
<td>16:30-18:00</td>
<td>ASSTA General Meeting</td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>SESSION A - Savoy West</td>
<td>SESSION B - Savoy 2</td>
</tr>
<tr>
<td>------------</td>
<td>--------------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>8:45-9:45</td>
<td>Plenary: Jane Stuart-Smith (University of Glasgow)</td>
<td>Twa son, some soldiers, and a city: An acoustic phonetic investigation of real-time change in Scottish English Savoy West</td>
</tr>
<tr>
<td>9:45-11:00</td>
<td>Morning Tea and Poster Session</td>
<td></td>
</tr>
<tr>
<td>11:00-11:25</td>
<td>NEW TECHNOLOGIES 1</td>
<td>A spectral analysis of laterals in three central Australian languages. Marija Tabian, Andy Butcher, Gavan Breen &amp; Richard Beare</td>
</tr>
<tr>
<td>11:25-11:50</td>
<td>Comparison of localised and feature-based variants of the mixelgram algorithm to perform audio visual speaker association. Trent Lewis &amp; Patrick Klaosen</td>
<td>Revisiting the pressure impulse in Australian languages: Bininj Gun-wok. Hywel Stoakes, Andy Butcher &amp; Janet Fletcher</td>
</tr>
<tr>
<td>11:50-12:15</td>
<td>Spectral enhancement of sounds by the stellate microcircuit of the ventral cochlear nucleus. Timothy Esler &amp; David Grayden</td>
<td>An acoustic study of the five Thai tones produced by ASD and TD children. Therdpong Thongseiratch, Jariya Chuthapisith &amp; Rungpat Roengpitya</td>
</tr>
<tr>
<td>12:15-12:40</td>
<td>Resources created for building New Zealand English voices. Catherine Watson &amp; Amelie Marchi</td>
<td>Mr White goes to market - running speech and citation tones in a Southern Thai bidialectal. Phil Rose</td>
</tr>
<tr>
<td>12:40-13:30</td>
<td>Lunch</td>
<td></td>
</tr>
<tr>
<td>13:30-14:30</td>
<td>Plenary: Gail Gillon (University of Canterbury)</td>
<td>Phonological awareness development in children with speech sound disorder Savoy West</td>
</tr>
<tr>
<td>14:30-14:55</td>
<td>ARTICULATORY PHONETICS</td>
<td>PERCEPTION 2</td>
</tr>
<tr>
<td>14:30-14:55</td>
<td>An electro-palatographic study of consonant sequences in Iwaidja. Janet Fletcher, Andrew Butcher &amp; Deborah Loakes</td>
<td>Improvements to vowel categorization in non-native regional accents resulting from multiple-versus single-talker training: A computational approach. Sarah Wright, Jason Shaw, Catherine Best, Gerard Docherty, Bronwen Evans, Paul Foulkes, Jen Hay &amp; Karen Mulak</td>
</tr>
<tr>
<td>15:20-15:45</td>
<td>Afternoon Tea</td>
<td></td>
</tr>
<tr>
<td>15:45-16:10</td>
<td>FORENSICS 2</td>
<td>SPEECH PRODUCTION 1</td>
</tr>
<tr>
<td>15:45-16:10</td>
<td>A first attempt at compensating for effects due to recording condition mismatch in formant trajectory based forensic voice comparison. Ewald Enzinger</td>
<td>Temporal planning in the production of Australian English compounds. Ivan Yuen, Nan Xu Rattanasone, Gretel McDonald, Rebecca Holt &amp; Katherine Demuth</td>
</tr>
<tr>
<td>16:10-16:35</td>
<td>Looking into real world: LR variability under forensically realistic conditions. Yuko Kinoshita</td>
<td>Time spent talking in retirement. Nina Fhärn, Frida Wigelius Skoglund &amp; Jan van Doorn</td>
</tr>
<tr>
<td>Time</td>
<td>Session A - Savoy West</td>
<td>Session B - Savoy 2</td>
</tr>
<tr>
<td>------------</td>
<td>-------------------------------------------------------------</td>
<td>----------------------------------------------------------</td>
</tr>
<tr>
<td>8:45-9:45</td>
<td>Plenary: T. Florian Jaeger (Rochester University)</td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Speech perception and adaptation as inference under uncertainty</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Savoy West</td>
<td></td>
</tr>
<tr>
<td>9:45-10:10</td>
<td>Morning Tea</td>
<td></td>
</tr>
<tr>
<td>10:10-10:35</td>
<td><strong>NEW TECHNOLOGIES 2</strong></td>
<td><strong>SPEECH PRODUCTION 2</strong></td>
</tr>
<tr>
<td></td>
<td>Automatic detection of speech truncation and speech rate.</td>
<td>Rentention of Spanish coda /s/ by speakers of Kashibo-</td>
</tr>
<tr>
<td></td>
<td>Chung Ting Justine Hui, Teh June Chin &amp; Catherine Watson</td>
<td>Kakataibo.</td>
</tr>
<tr>
<td>10:35-11:00</td>
<td>Short utterance PLDA speaker verification using SN-WLDA</td>
<td>Assibilition in Trans-New Guinea languages of the Bird's</td>
</tr>
<tr>
<td></td>
<td>and variance modelling techniques.</td>
<td>Head region.</td>
</tr>
<tr>
<td></td>
<td>Ahilan Kanagasundaram, David Dean &amp; Sridha Sridharan</td>
<td>Fanny Cottet</td>
</tr>
<tr>
<td>11:00-11:25</td>
<td>**Recalling clustering trees using impact ratios for robust</td>
<td><strong>Iconicity in Korean consonantal symbolism.</strong></td>
</tr>
<tr>
<td></td>
<td>hierarchical speaker clustering.</td>
<td>Nahyun Kwon</td>
</tr>
<tr>
<td></td>
<td>Houman Ghaemmaghami, David Dean, Shahram</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Kalantari &amp; Sridha Sridharan</td>
<td></td>
</tr>
<tr>
<td>11:25-11:50</td>
<td>Phonetic spoken term search using topic information.</td>
<td>Some initial findings regarding first language influence</td>
</tr>
<tr>
<td></td>
<td>Sharam Kalantari, David Dean &amp; Sridha Sridharan</td>
<td>on brass playing instruments.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Matthias Heyne &amp; Donald Derrick</td>
</tr>
<tr>
<td>11:50-12:40</td>
<td>Lunch</td>
<td></td>
</tr>
<tr>
<td>12:40-13:05</td>
<td><strong>PROSODY</strong></td>
<td><strong>PERCEPTION 3</strong></td>
</tr>
<tr>
<td></td>
<td>Sociophonetic and prosodic influences on judgements of</td>
<td>A restriction of minimal words, or word endings?</td>
</tr>
<tr>
<td></td>
<td>sentence type.</td>
<td>Jason Brown &amp; Forrest Panther</td>
</tr>
<tr>
<td></td>
<td>Paul Warren</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Casey Tait, Marija Tabian &amp; Ingrid Skyes</td>
<td>Grant McGuire</td>
</tr>
<tr>
<td>13:30-13:55</td>
<td><strong>Examining the influence of pitch accents on word learning</strong></td>
<td>Frequency in the input affects perception of phonological</td>
</tr>
<tr>
<td></td>
<td>Michael Walsh, Katrin Schweitzer, Hinrich Schütze &amp;</td>
<td>contrasts for native speakers.</td>
</tr>
<tr>
<td></td>
<td>Dermot Lynott</td>
<td>Rikke Louise Bundgaard-Neilsen &amp; Brett Baker</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13:55-14:20</td>
<td>Afternoon Tea</td>
<td></td>
</tr>
<tr>
<td>14:20-14:45</td>
<td><strong>PERCEPTION 4</strong></td>
<td><strong>CHILD LANGUAGE ACQUISITION</strong></td>
</tr>
<tr>
<td></td>
<td>Influence of phonological, morphological, and prosodic</td>
<td>Consonant inventory of infants aged 0-6 months.</td>
</tr>
<tr>
<td></td>
<td>factors on phoneme detection by native and second-</td>
<td>Adele Gregory, Marija Tabain &amp; Michael Robb</td>
</tr>
<tr>
<td></td>
<td>language adults.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Valeria Peretokina, Michael Tyler and Catherine Best</td>
<td></td>
</tr>
<tr>
<td>14:45-15:10</td>
<td>Perception of Italian and Japanese consonant length by</td>
<td>Investigating the effect of intrusive noise levels on</td>
</tr>
<tr>
<td></td>
<td>native speakers of Australian English and Italian: A pilot</td>
<td>speech perception in an open-plan kindergarten classroom.</td>
</tr>
<tr>
<td></td>
<td>study.</td>
<td>Kiri Mealings, Katherine Demuth, Jorg Buchholz &amp; Harvey</td>
</tr>
<tr>
<td></td>
<td>Kimiko Tsukada, Felicity Cox, John Hajek &amp; Yukari Hirata</td>
<td>Dillon</td>
</tr>
<tr>
<td>15:10-15:35</td>
<td><strong>Is more always better? The perception of Dutch vowels by</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>English versus Spanish listeners.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Samra Alispahic, Paola Escudero &amp; Karen Mulak</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>FINISH</td>
<td></td>
</tr>
</tbody>
</table>
# Table of Contents for Papers

## Forensics 1
10:25-12:05, 3rd December 2014  
Session Chair: Viktoria Papp

<table>
<thead>
<tr>
<th>No.</th>
<th>Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>LR-based forensic comparison under severe test-data scarcity.</td>
<td>Yuko Kinoshita¹, Michael Wagner¹</td>
</tr>
<tr>
<td></td>
<td></td>
<td>¹University of Canberra</td>
</tr>
<tr>
<td>20</td>
<td>Replicate mismatch between test and background/development databases: The effect on the performance of likelihood ratio-based forensic voice comparison</td>
<td>Shunichi Ishihara³</td>
</tr>
<tr>
<td></td>
<td></td>
<td>³Australian National University</td>
</tr>
<tr>
<td>24</td>
<td>Regional variation and the definition of the relevant population in likelihood ratio-based forensic voice comparison using cepstral coefficients</td>
<td>Vincent Hughes⁷, Paul Foulkes¹</td>
</tr>
<tr>
<td></td>
<td></td>
<td>¹University of York</td>
</tr>
<tr>
<td>28</td>
<td>Comparison between speech parameters for forensic voice comparison using mobile phone speech</td>
<td>Esam Alzqhol¹, Balamurali B T Nair³, Bernard Guillemin¹</td>
</tr>
<tr>
<td></td>
<td></td>
<td>¹University of Auckland</td>
</tr>
</tbody>
</table>

## PANZE
10:25-12:05, 3rd December 2014  
Session Chair: Kevin Watson

<table>
<thead>
<tr>
<th>No.</th>
<th>Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>32</td>
<td>Phonologisation of vowel duration and nasalised /æ/ in Australian English</td>
<td>Felicity Cox¹, Sallianne Palethorpe²</td>
</tr>
<tr>
<td></td>
<td></td>
<td>¹Macquarie University, ²ARC Centre of Excellence in Cognition and its Disorders</td>
</tr>
<tr>
<td>36</td>
<td>Why the SQUARE vowel is the most variable in Sydney</td>
<td>Nhung Nguyen¹, Jason Shaw¹</td>
</tr>
<tr>
<td></td>
<td></td>
<td>¹University of Western Sydney</td>
</tr>
<tr>
<td>40</td>
<td>Identifying /æl/-/æl/: A comparison between two regional Australian towns</td>
<td>Deborah Loakes¹, John Hajek¹, Joshua Clothier¹, Janet Fletcher¹</td>
</tr>
<tr>
<td></td>
<td></td>
<td>¹University of Melbourne</td>
</tr>
<tr>
<td>44</td>
<td>Comparing acoustic analyses of Australian English vowels from Sydney: Cox (2006) versus AusTalk</td>
<td>Jaydene Elvin¹, Paola Escudero¹</td>
</tr>
<tr>
<td></td>
<td></td>
<td>¹University of Western Sydney</td>
</tr>
</tbody>
</table>

## L2 Acquisition
14:00-14:50, 3rd December 2014  
Session Chair: Jeanette King

<table>
<thead>
<tr>
<th>No.</th>
<th>Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>49</td>
<td>Segmental and tonal errors in L2 Mandarin speech produced by Australian English learners</td>
<td>Wentao Gu¹, Ting Zhang¹, Chiharu Tsurutani²</td>
</tr>
<tr>
<td></td>
<td></td>
<td>¹Nanjing Normal University, ²Griffith University</td>
</tr>
<tr>
<td>53</td>
<td>Does immersion experience reduce /r/-/l/ category overlap for Japanese learners of English?</td>
<td>Michael Tyler¹, Saya Kawase¹, Mark Antoniou¹</td>
</tr>
<tr>
<td></td>
<td></td>
<td>¹University of Western Sydney</td>
</tr>
</tbody>
</table>

## Pitch
15:15-16:30, 3rd December 2014  
Session Chair: James Gruber

<table>
<thead>
<tr>
<th>No.</th>
<th>Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>58</td>
<td>Tonal Alignment of Focal Pitch Accents in Two Varieties of Indian English</td>
<td>Olga Maxwell¹, Janet Fletcher¹</td>
</tr>
<tr>
<td></td>
<td></td>
<td>¹University of Melbourne</td>
</tr>
<tr>
<td>62</td>
<td>Amplitude and F0 as acoustic correlates of Kelantan Malay word-initial geminates</td>
<td>Mohd Hilmi Hamzah¹, Janet Fletcher², John Hajek²,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>¹Universiti Teknologi Malaysia, ²University of Melbourne</td>
</tr>
<tr>
<td>66</td>
<td>Pitch accents and prosodic properties of the Clitic in Yukulta (Tangkic)</td>
<td>Cicely Bonnin¹</td>
</tr>
<tr>
<td></td>
<td></td>
<td>¹University of Queensland</td>
</tr>
<tr>
<td>Session</td>
<td>Title</td>
<td>Authors</td>
</tr>
<tr>
<td>---------</td>
<td>----------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------</td>
</tr>
<tr>
<td>71</td>
<td>Effects of short-term exposure to unfamiliar regional accents: Australians' categorization of London and Yorkshire English consonants</td>
<td>Jason Shaw¹, Catherine Best², Karen Mulak³, Gerard Docherty², Bronwen Evans³, Paul Foulkes⁴, Jen Hay⁵, Jalal Al-Tamimi⁶, Mike Peek⁷, Katharine Mair⁸, Sophie Wood⁹</td>
</tr>
<tr>
<td>75</td>
<td>Listeners cope with speaker and accent variation differently: Evidence from the Go/No-go task</td>
<td>Buddhahas Kriengwatana¹, Paola Escudero¹, Josephine Terry²</td>
</tr>
<tr>
<td>79</td>
<td>Acoustic distance explains speaker versus accent normalization in infancy</td>
<td>Paola Escudero¹, Karen Mulak¹, Samra Alispahic¹</td>
</tr>
<tr>
<td>84</td>
<td>Comparison of localised and feature-based variants of the mixelgram algorithm to perform audio-visual speaker association</td>
<td>Trent Lewis¹, Patrick Klaosen¹</td>
</tr>
<tr>
<td>88</td>
<td>Spectral enhancement of sounds by the stellate microcircuit of the ventral cochlear nucleus</td>
<td>Timothy Esler¹, David Grayden¹</td>
</tr>
<tr>
<td>92</td>
<td>Resources created for building New Zealand English voices</td>
<td>Catherine Watson¹, Amelie Marchi²</td>
</tr>
<tr>
<td>97</td>
<td>A spectral analysis of laterals in three Central Australian languages</td>
<td>Marija Tabain¹, Andy Butcher¹, Gavan Breen¹, Richard Beare¹</td>
</tr>
<tr>
<td>101</td>
<td>Revisiting the pressure impulse in Australian languages: Bininj Gun-wok</td>
<td>Hywel Stoakes¹, Andy Butcher², Janet Fletcher²</td>
</tr>
<tr>
<td>105</td>
<td>An acoustic study of the five Thai tones produced by ASD and TD children</td>
<td>Therdpong Thongseiratch¹, Jariya Chuthapisith¹, Rungpat Roengpitya¹</td>
</tr>
<tr>
<td>109</td>
<td>Mr. White goes to market - running speech and citation tones in a southern Thai bidialectal</td>
<td>Phil Rose¹,²</td>
</tr>
<tr>
<td>114</td>
<td>An electro-palatographic study of consonant sequences in Iwaidja</td>
<td>Janet Fletcher¹, Andrew Butcher², Deborah Loakes¹</td>
</tr>
<tr>
<td>118</td>
<td>’Advanced Tongue Root’ in Lopit: Acoustic and ultrasound evidence</td>
<td>Rosey Billington¹</td>
</tr>
<tr>
<td>Session</td>
<td>Time</td>
<td>Chair</td>
</tr>
<tr>
<td>----------</td>
<td>-----------------------</td>
<td>------------------------</td>
</tr>
<tr>
<td>Perception 2</td>
<td>14:30-15:20, 4&lt;sup&gt;th&lt;/sup&gt; December 2014</td>
<td>Lynn Clark</td>
</tr>
<tr>
<td>Thai phonetically balanced word recognition test: Test-retest reliability and error analysis</td>
<td>A. Munthuli&lt;sup&gt;1&lt;/sup&gt;, C. Tantibundhit&lt;sup&gt;1&lt;/sup&gt;, C. Onsuwan&lt;sup&gt;1&lt;/sup&gt;, K. Kosawat&lt;sup&gt;2&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Forensics 2</td>
<td>15:45-16:10, 4&lt;sup&gt;th&lt;/sup&gt; December 2014</td>
<td>Paul Foulkes</td>
</tr>
<tr>
<td>Looking into the real world: LR variability under forensically realistic conditions</td>
<td>Yuko Kinoshita&lt;sup&gt;1&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Speech Production 1</td>
<td>14:30-15:20, 4&lt;sup&gt;th&lt;/sup&gt; December 2014</td>
<td>Catherine Theys</td>
</tr>
<tr>
<td>Time spent talking in retirement</td>
<td>Nina Fhärm&lt;sup&gt;1&lt;/sup&gt;, Frida Wigelius Skoglund&lt;sup&gt;1&lt;/sup&gt;, Jan van Doorn&lt;sup&gt;1&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>New Technologies 2</td>
<td>10:10-11:50 5&lt;sup&gt;th&lt;/sup&gt; December 2014</td>
<td>Catherine Watson</td>
</tr>
<tr>
<td>Short utterance PLDA speaker verification using SN-WLDA and variance modelling techniques</td>
<td>Ahilan Kanagasundaram&lt;sup&gt;1&lt;/sup&gt;, David Dean&lt;sup&gt;1&lt;/sup&gt;, Sridha Sridharan&lt;sup&gt;1&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Rescaling clustering trees using impact ratios for robust hierarchical speaker clustering</td>
<td>Houman Ghaemmaghami&lt;sup&gt;1&lt;/sup&gt;, David Dean&lt;sup&gt;1&lt;/sup&gt;, Shahram Kalantari&lt;sup&gt;1&lt;/sup&gt;, Sridha Sridharan&lt;sup&gt;1&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Phonetic spoken term search using topic information</td>
<td>Shahram Kalantari&lt;sup&gt;1&lt;/sup&gt;, David Dean&lt;sup&gt;1&lt;/sup&gt;, Sridha Sridharan&lt;sup&gt;1&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Speech Production 2</td>
<td>10:10-11:50 5&lt;sup&gt;th&lt;/sup&gt; December 2014</td>
<td>Beth Hume</td>
</tr>
<tr>
<td>Assibilation in Trans-New Guinea languages of the Bird’s Head region</td>
<td>Fanny Cottet&lt;sup&gt;1&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Page</td>
<td>Title</td>
<td>Authors</td>
</tr>
<tr>
<td>------</td>
<td>-----------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>175</td>
<td>Iconicity in Korean consonantal symbolism</td>
<td>Nahyun Kwon¹</td>
</tr>
<tr>
<td></td>
<td></td>
<td>¹University of Queensland</td>
</tr>
<tr>
<td>179</td>
<td>Some initial findings regarding first language influence on playing brass instruments</td>
<td>Matthias Heyne¹, Donald Derrick¹,²</td>
</tr>
<tr>
<td></td>
<td></td>
<td>¹University of Canterbury, ²University of Western Sydney</td>
</tr>
<tr>
<td></td>
<td><strong>Prosody</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>12:40-13:55, 5th December 2014</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Session Chair: James Gruber</strong></td>
<td></td>
</tr>
<tr>
<td>184</td>
<td>Sociophonic and prosodic influences on judgements of sentence type</td>
<td>Paul Warren¹</td>
</tr>
<tr>
<td></td>
<td></td>
<td>¹Victoria University of Wellington</td>
</tr>
<tr>
<td>188</td>
<td>Stress-meter alignment in American hip hop</td>
<td>Casey Tait¹, Marija Tabain¹, Ingrid Sykes¹</td>
</tr>
<tr>
<td></td>
<td></td>
<td>¹La Trobe University</td>
</tr>
<tr>
<td>192</td>
<td>Examining the influence of pitch accents on word learning in German</td>
<td>Michael Walsh¹, Katrin Schweitzer¹, Hinrich Schütze¹, Dermot Lynott¹</td>
</tr>
<tr>
<td></td>
<td></td>
<td>¹University of Stuttgart, ¹Ludwig-Maximilians University Munich, ³Lancaster University</td>
</tr>
<tr>
<td></td>
<td><strong>Perception 3</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>12:40-13:55, 5th December 2014</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Session Chair: Michael Tyler</strong></td>
<td></td>
</tr>
<tr>
<td>197</td>
<td>A restriction on minimal words, or word endings?</td>
<td>Jason Brown¹, Forrest Panther¹</td>
</tr>
<tr>
<td></td>
<td></td>
<td>¹University of Auckland</td>
</tr>
<tr>
<td>201</td>
<td>Orthographic effects on phonetic cue weighting</td>
<td>Grant McGuire¹</td>
</tr>
<tr>
<td></td>
<td></td>
<td>¹University of California Santa Cruz</td>
</tr>
<tr>
<td>205</td>
<td>Frequency in the input affects perception of phonological contrasts for native speakers</td>
<td>Rikke Louise Bundgaard-Nielsen¹,², Brett Baker³</td>
</tr>
<tr>
<td></td>
<td></td>
<td>¹La Trobe University, ¹University of Western Sydney, ³University of Melbourne</td>
</tr>
<tr>
<td></td>
<td><strong>Perception 4</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>14:20-15:35 5th December 2014</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Session Chair: Jen Hay</strong></td>
<td></td>
</tr>
<tr>
<td>210</td>
<td>Influence of phonological, morphological, and prosodic factors on phoneme detection by native and second-language adults</td>
<td>Valeria Peretokina¹, Michael Tyler¹, Catherine Best¹</td>
</tr>
<tr>
<td></td>
<td></td>
<td>¹University of Western Sydney</td>
</tr>
<tr>
<td>214</td>
<td>Perception of Italian and Japanese consonant length by native speakers of Australian English and Italian: A pilot study</td>
<td>Kimiko Tsukada¹, Felicity Cox¹, John Hajek¹, Yukari Hirata¹</td>
</tr>
<tr>
<td></td>
<td></td>
<td>¹Macquarie University, ¹University of Melbourne, ²Colgate University</td>
</tr>
<tr>
<td>218</td>
<td>Is more always better? The perception of Dutch vowels by English versus Spanish listeners</td>
<td>Samra Alispahic¹, Paola Escudero¹, Karen Mulak¹</td>
</tr>
<tr>
<td></td>
<td></td>
<td>¹University of Western Sydney</td>
</tr>
<tr>
<td></td>
<td><strong>Child Language Acquisition</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>14:20-15:10 5th December 2014</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Session Chair: Jeanette King</strong></td>
<td></td>
</tr>
<tr>
<td>223</td>
<td>Consonant inventory of infants aged 0-6 months</td>
<td>Adele Gregory¹, Marija Tabain¹, Michael Robb¹</td>
</tr>
<tr>
<td></td>
<td></td>
<td>¹La Trobe University, ¹University of Canterbury</td>
</tr>
<tr>
<td>227</td>
<td>Investigating the effect of intrusive noise levels on speech perception in an open-plan kindergarten classroom</td>
<td>Kiri Mealings¹, Katherine Demuth¹, Jorg Buchholz¹,², Harvey Dillon³</td>
</tr>
<tr>
<td></td>
<td></td>
<td>¹Macquarie University, ¹National Acoustics Laboratories</td>
</tr>
<tr>
<td>Posters</td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>242</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>The voiced emphatic coronal stop [ḍ] in the Ḥaḍrami Arabic: A socio-phonetic study</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hamad Altari¹</td>
<td></td>
<td></td>
</tr>
<tr>
<td>¹University of Auckland</td>
<td></td>
<td></td>
</tr>
<tr>
<td>243</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Contour-based analysis of EGG data from words in isolation and connected speech</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stephen Bier³, Catherine Watson³, Clare McCann³</td>
<td></td>
<td></td>
</tr>
<tr>
<td>³University of Auckland</td>
<td></td>
<td></td>
</tr>
<tr>
<td>244</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Venezuelan Spanish intransitives: More prosodic than you’d think</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sasha Calhoun¹, Erwin La Cruz¹, Ana Olssen¹</td>
<td></td>
<td></td>
</tr>
<tr>
<td>¹Victoria University of Wellington</td>
<td></td>
<td></td>
</tr>
<tr>
<td>245</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Recency effects on word-medial /t/ in New Zealand English: initial observations</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lynn Clark³, Liam Walsh¹</td>
<td></td>
<td></td>
</tr>
<tr>
<td>¹University of Canterbury</td>
<td></td>
<td></td>
</tr>
<tr>
<td>246</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>F0 as a word boundary cue for segmenting New Zealand English ethnolects</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kylie Fitzgerald¹, Viktoria Papp¹, Jen Hay⁴</td>
<td></td>
<td></td>
</tr>
<tr>
<td>¹University of Canterbury</td>
<td></td>
<td></td>
</tr>
<tr>
<td>247</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Distinguishing dysarthric speech: Vowel acoustics and measurements</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Annalise Fletcher¹, Megan McAuliffe³, Kaitlin Lansford³, Julie Liss³</td>
<td></td>
<td></td>
</tr>
<tr>
<td>¹University of Canterbury, ³Florida State University, ³Arizona State University</td>
<td></td>
<td></td>
</tr>
<tr>
<td>248</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Ability to identify unfamiliar speech sounds negatively correlates with second language proficiency</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yurika Hashimoto¹, Ian Wilson³, Younghyon Heo¹</td>
<td></td>
<td></td>
</tr>
<tr>
<td>¹University of Aizu</td>
<td></td>
<td></td>
</tr>
<tr>
<td>249</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Why Indo-Aryan languages adapt English alveolars as retroflexes: Acoustic evidence from Punjabi</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Qandeel Hussain¹</td>
<td></td>
<td></td>
</tr>
<tr>
<td>¹Macquarie University</td>
<td></td>
<td></td>
</tr>
<tr>
<td>250</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>I Bag Your Pardon: The Albertan æ/ɛ shift and community grammars</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jacqueline Jones¹, Stephen Winters¹</td>
<td></td>
<td></td>
</tr>
<tr>
<td>¹University of Calgary</td>
<td></td>
<td></td>
</tr>
<tr>
<td>251</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Diphthong trajectories in Māori</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jeanette King¹, Catherine Watson³, Margaret MacLagan¹, Peter Keegan², Ray Harlow³</td>
<td></td>
<td></td>
</tr>
<tr>
<td>¹University of Canterbury, ²University of Auckland, ³University of Waikato (Retired)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>253</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Preliminary comparison of New Caledonian &amp; Metropolitan French mid vowels</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eleanor Lewis¹</td>
<td></td>
<td></td>
</tr>
<tr>
<td>¹University of Melbourne</td>
<td></td>
<td></td>
</tr>
<tr>
<td>254</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Eye movements reveal Cantonese listeners use statistical information to assess category membership of acoustic cues</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jessie Nixon³, Jacolien van Rij³, Harald Baayen³, Peggy Mok², Yiya Chen³</td>
<td></td>
<td></td>
</tr>
<tr>
<td>²University of Tübingen, ³Chinese University of Hong Kong, ³Leiden University</td>
<td></td>
<td></td>
</tr>
<tr>
<td>255</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Articulatory coordination in Nama click consonants</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Michael Prouctor¹, Yinghua Zhu², Adam Lammert², Asterios Toutios², Bonny Sands³, Shrikanth Narayanan³</td>
<td></td>
<td></td>
</tr>
<tr>
<td>¹Macquarie University, ²University of Southern California, ³Northern Arizona University</td>
<td></td>
<td></td>
</tr>
<tr>
<td>256</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Blocking in linguistic associative learning</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Darcy Rose¹, Peter Racz¹, Jen Hay¹, Beth Hume¹</td>
<td></td>
<td></td>
</tr>
<tr>
<td>University of Canterbury</td>
<td></td>
<td></td>
</tr>
<tr>
<td>257</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Synthesizing speech using the AusTalk corpus</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zhijie Shao⁵, Richard Leibbrandt⁵, Trent Lewis⁵</td>
<td></td>
<td></td>
</tr>
<tr>
<td>⁵Flinders University</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Page</td>
<td>Title</td>
<td>Authors</td>
</tr>
<tr>
<td>------</td>
<td>----------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------------</td>
</tr>
</tbody>
</table>
| 258  | The neurophysiological correlates of overt speech productions: Dissociation of the speech production phases and test-retest reliability | Catherine Theys¹, Maarten De Vos², Megan McAuliffe¹  
¹University of Canterbury, ²University of Oldenbury |                                                  |
| 259  | Intergenerational vowel change in several Russian-English speakers    | Ben Volchok¹  
¹University of Melbourne                                                             |                                                  |