Eliciting natural conversational Lombard speech in realistic acoustic environments

Timothy Beechey[1,2,3], Jörg Buchholz[1,2,3] & Gitte Keidser[1,3]

1. National Acoustic Laboratories 2. Macquarie University 3. The HEARing CRC

Background

- Speech intelligibility tests, which are commonly used in the audiology clinic and in research, do not reflect what people experience outside the clinic.
- Speech intelligibility tests may be poor predictors of communication ability and effort in daily life (Working Group on Speech Understanding and Aging, 1988).
- An alternative approach is to measure communication difficulty and effort from speech production which also incorporates more realistic task demands.

Objectives

This study aimed to develop and verify a framework for assessing communication difficulty and effort using a conversation paradigm. Our objectives were to:

- Develop and verify a conversation elicitation task to elicit fluent, balanced conversational speech and language.
- Determine a set of measures which are sensitive to changes in communication difficulty across acoustic environments.
- Establish baseline data for normal hearers before testing people with hearing impairment in the future.
- Consider higher level linguistic effects in addition to the acoustic and phonetic changes considered in Lombard studies (Cooke & Lu, 2010; Lane & Tranel, 1971)

Hypotheses

- Talkers will modify their speech and language in line with the difficulty of the acoustic environment (rather than a two-way quiet/noise distinction)
- Changes at acoustic and phonetic levels will occur sooner (in less complex acoustic environments) than changes at higher linguistic levels

Variables of interest

- **Acoustic** Vocal level (broadband and gammatone bands)
- **Acoustic-phonetic** Formant frequencies, formant bandwidths, F0, vowel duration
- **Lexical and utterance level** Utterance duration, pause duration

Methods

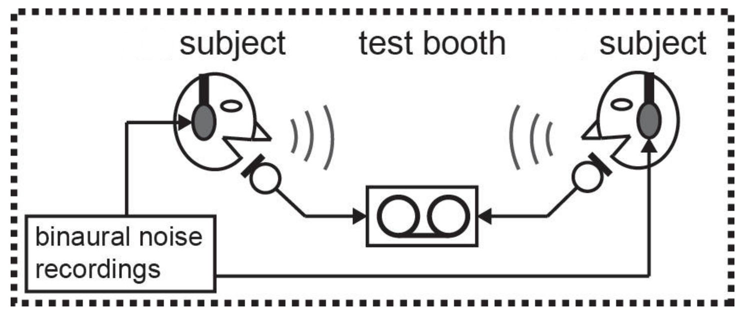
Participants 20 native Australian English speakers aged between 18 and 40 with normal puretone hearing thresholds including 10 males and 10 females were tested in pairs.

Acoustic environments 5 acoustic scenes recorded using a 62 channel hard-sphere microphone array were presented at their original levels. All participants completed subjective rating scales regarding the difficulty of conversation in each environment.

Acoustic environments and levels (dB A)

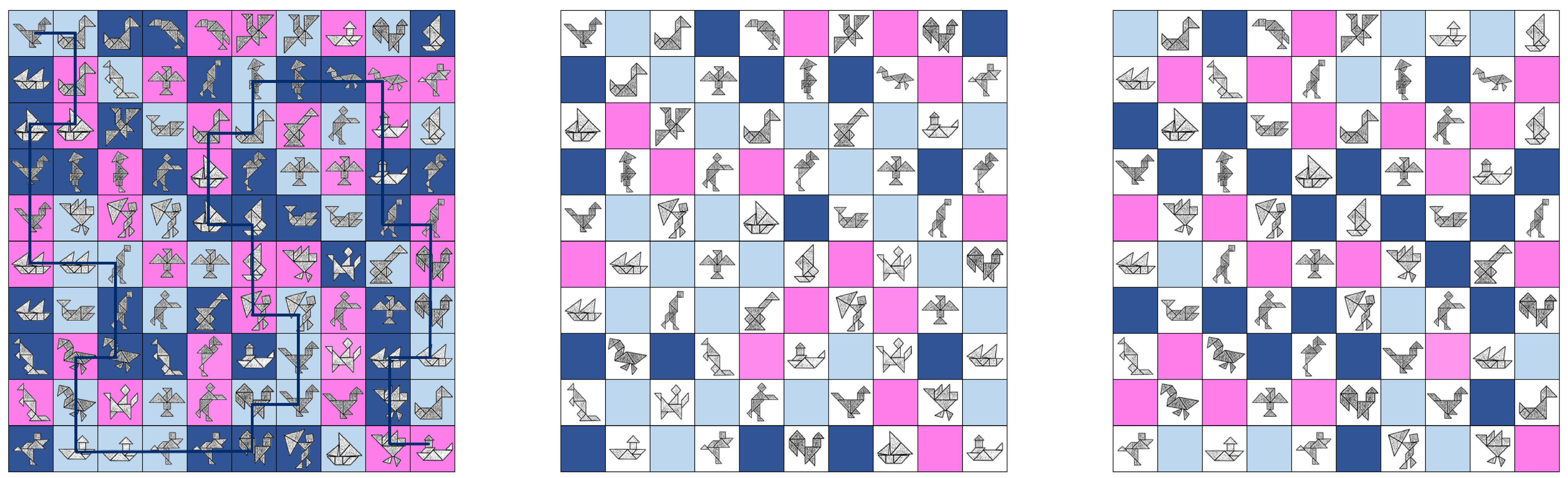
Library	Office	Cafe	Traffic	Foodcourt
48.5	56.5	76.4	79.7	81.8

Experiment set-up



Puzzle task Participants completed puzzles where the task was to find a path by moving between matching colors or matching pictures. Each participant saw a complementary version of the same puzzle. Tangram symbols were used to elicit complex descriptions and questions.

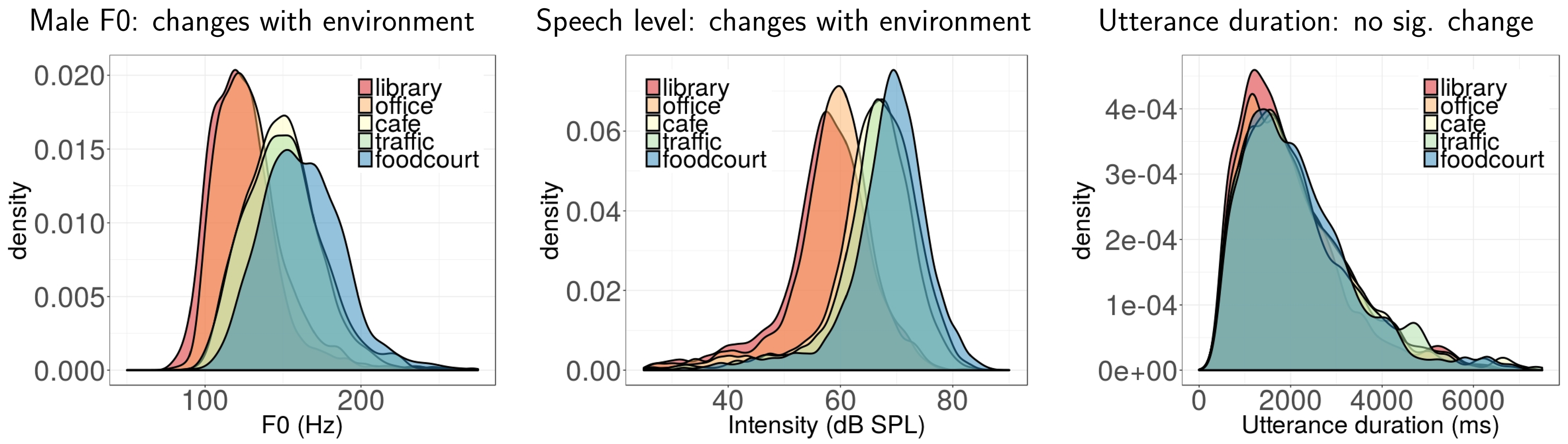
Complete (unseen) puzzle — Participant 1 view — Participant 2 view



Procedure Pairs of participants sat facing one-another at a distance of 1.3 m. Each noise was presented through open headphones (Beyerdynamic DT 990 Pro) in 5 minute blocks in counter-balanced order. Speech was recorded using close-talk microphones which were calibrated to a center microphone at 1m distance.

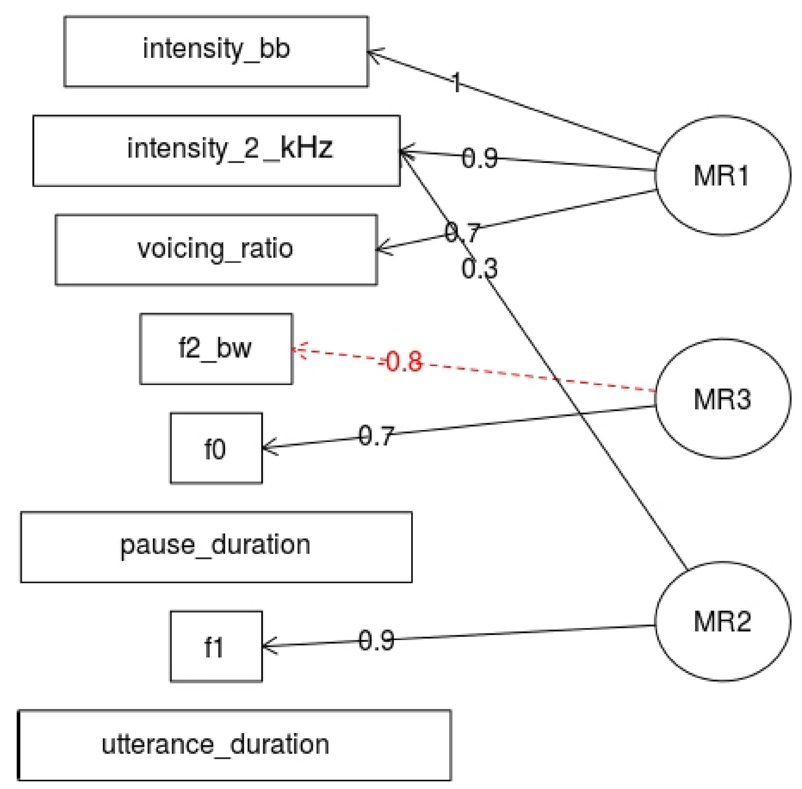
Results

Analyses indicated significant differences for most acoustic and phonetic variables but no significant changes across environments at higher linguistic levels. See, for example, utterance duration below:



Factor analysis

Exploratory factor analysis suggested the following latent structure:



Confirmatory factor analysis returned a good fit with utterance duration and pause duration included in MR2 .

χ^2	χ^2/df	p-value	CFI	RMSE	SRMR	MFI
25.59	1.59	.06	.985	.077	.06	.953

Discussion & future research

- The conversation task successfully elicited natural conversational speech and language.
- Changes at acoustic and phonetic levels related to changes in complexity of environments were observed for the normal hearing participants.
- A future study will investigate whether changes occur at higher linguistic levels in conversations involving hearing impaired participants.

References

Cooke, M., & Lu, Y. (2010). Spectral and temporal changes to speech produced in the presence of energetic and informational maskers. *JASA*, *128*(4), 2059–2069.

Lane, H., & Tranel, B. (1971). The Lombard sign and the role of hearing in speech. *JSLRH*, *14*(4), 677–709.

Working Group on Speech Understanding and Aging. (1988). Speech understanding and aging. *JASA*(83), 859–895.

Acknowledgements

This project was supported by funding from the HEARing CRC and the Oticon Foundation.

tim.beechey@nal.gov.au