

Optimising Wrist-Worn Haptic Hearing Aids for Individual Needs: Targeted Audio-to-Tactile Conversion Improves Fricative Perception

Siu Ching Tse^{3, 4}, Mark Fletcher^{1, 2, 3} & Carl Verschuur^{1, 3}

¹ Faculty of Medicine, ² University of Southampton Auditory Implant Service, ³ Institute of Sound and Vibration Research, University of Southampton, University Road, Southampton, SO17 1BJ, United Kingdom
⁴ Norfolk and Norwich University Hospitals NHS Foundation Trust, Colney Lane, Norwich, NR4 7UY, United Kingdom



Background



Natural speech

Audio-to-Tactile Conversion
via a haptic hearing aid

Haptic signals conveying speech
information through the skin

Cochlear Implants (CI) for speech:

- ✗ Some fine acoustic details for speech
- Less info. for place of articulation
- Substantial individual variations
- Poorer listeners show vowel & consonant perception errors

Electro-Haptic Stimulation:

- EHS: Haptic stimulation augments electrical signals
- Compact actuators + batteries + microprocessors = lower-cost, readily available, safe & wearable
- Benefits: phoneme discrimination, phrase & word identification, speech perception in noise, spatial hearing & sound localisation

Figure 1. Principle of Electro-Haptic Stimulation.

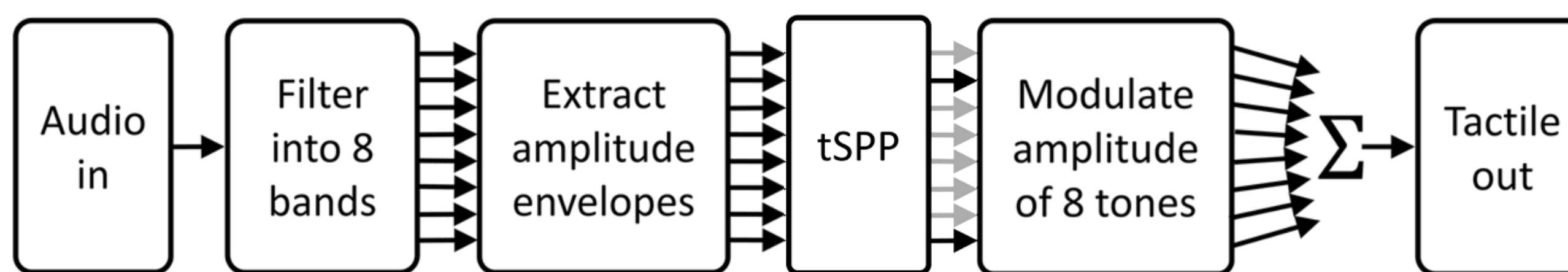


Figure 2. A block diagram of the stages of audio-to-tactile vocoding for signal conversion. Extracted amplitude envelopes are subjected to tactile spectral peak picking (tSPP) and used to modulate the amplitudes of 8 vibrotactile tones. The combined tones form the output signal.

MosaicTwo haptic hearing aid is manufactured by 3D printing. It consists of a central signal processor, a low-frequency motor and a high-frequency motor.



Previous research¹:

- **MosaicTwo**: wrist-worn haptic hearing aid
- **“Formant Focusing”** algorithm:
 - ✓ Tactile vowel discrimination
 - ✗ Optimised algorithms for speech
 - ✗ Effective algorithms for plosives & fricatives
 - ? Are targeted, phoneme-specific conversion algorithms using discontinuous filter bands effective?

Research Aim:

- To create 3 new signal conversion algorithms based on actual phonemic challenges
- To compare the 3 new algorithms with the existing Formant Focusing (baseline)
- To evaluate the effectiveness of all algorithms in enhancing tactile phoneme discrimination
- To verify the effectiveness of phoneme-specific algorithms for a specific class of phonemes

Methodology

Designing 3 new algorithms

Target patient groups

- Poorer CI listeners
- HL without HA

Phoneme error patterns

- Reinterpreting existing empirical data

Characteristic frequencies of challenging phonemes

- Literature review

Mapping characteristic frequencies onto filter band frequencies

- Selecting filter band frequency ranges to generate tactile output

Audio speech stimuli

- EHS Research Group Phoneme Corpus
- 44 British English phonemes
- 2 Male & 2 Female voice recordings
- 45 pairs of phoneme contrasts for the experimental task

Participants

- N=30, aged 23-40 (mean = 26.8)
- 11 male, 19 female
- 17 native English speakers
- Normal tactile sensation

Procedures

- Temperature-controlled, vibration-isolated room
- MosaicTwo @ left wrist
- 55 dB(A) pink noise masker

Screening phase:

- All passed with >90% accuracy
- ≥10 dB dynamic range in tactile sensitivity across all tested frequencies at the left wrist

Experimental phase:

- **Three-interval, three-alternative forced-choice phoneme discrimination task**
- 3 vibrotactile stimuli: “odd one out”
- **720 stimuli** in 2 blocks of 360 trials
- 45 phoneme pairs x 4 algorithms x 4 recordings
- **Automatic randomisation** for stimuli
- Can repeat each trial (< 8 times)
- **%correct** for tactile phoneme discrimination

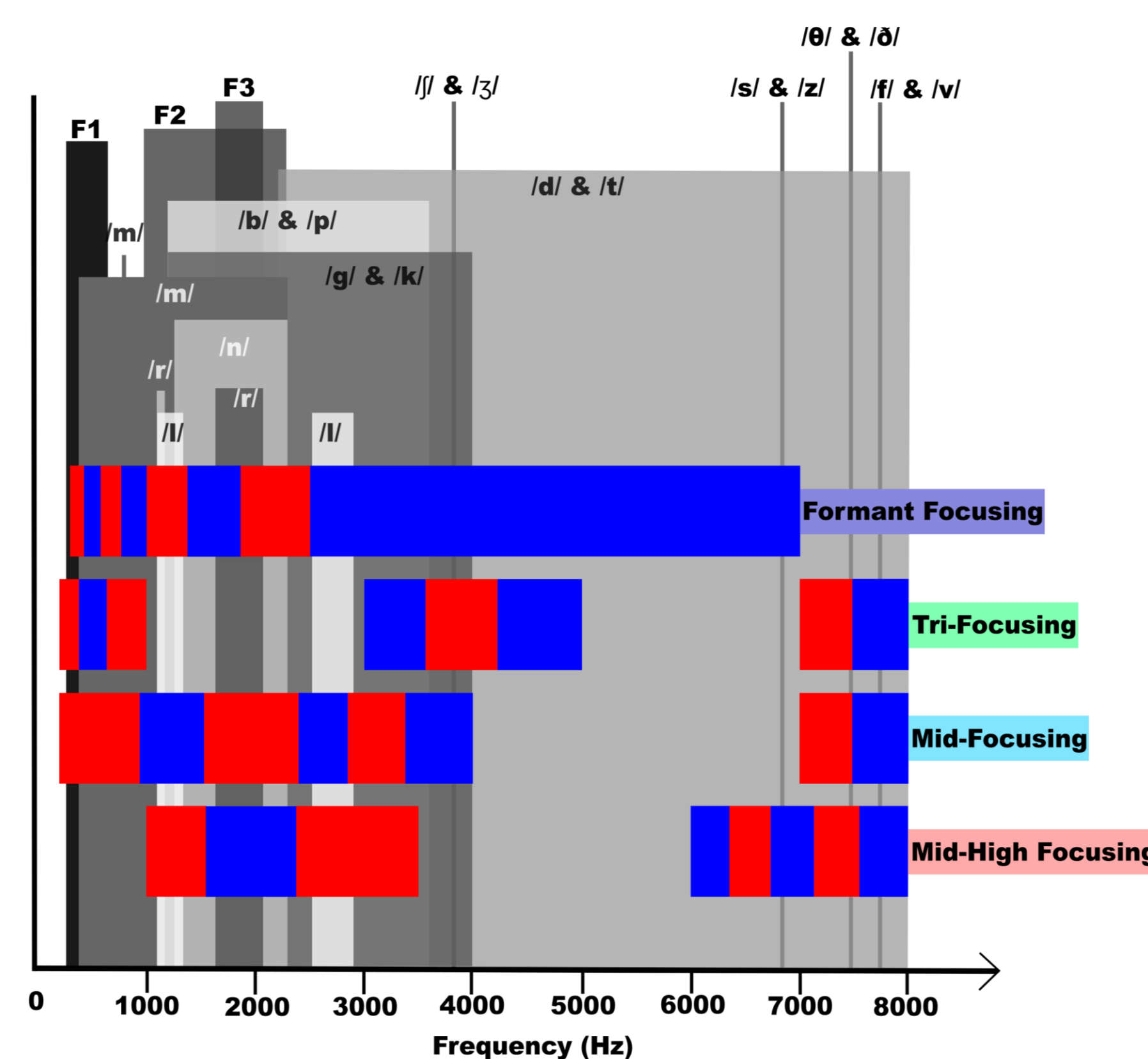


Figure 3. Filter band configurations of the four algorithms, mapped against the characteristic frequencies of the phoneme targets.

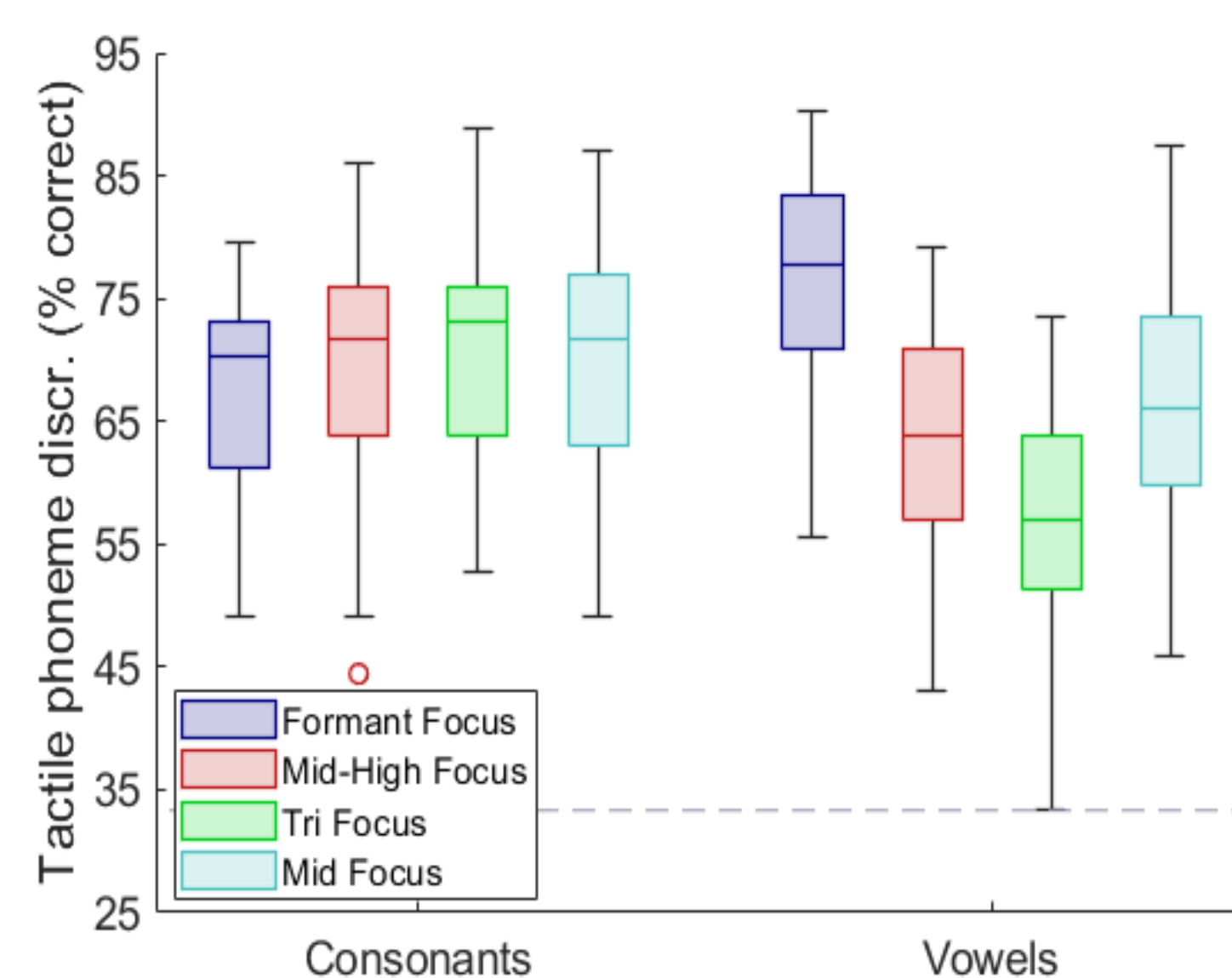


Figure 4. Box-and-whisker plot showing %correct for all algorithms across 2 phoneme classes. Chance level at 33.3%

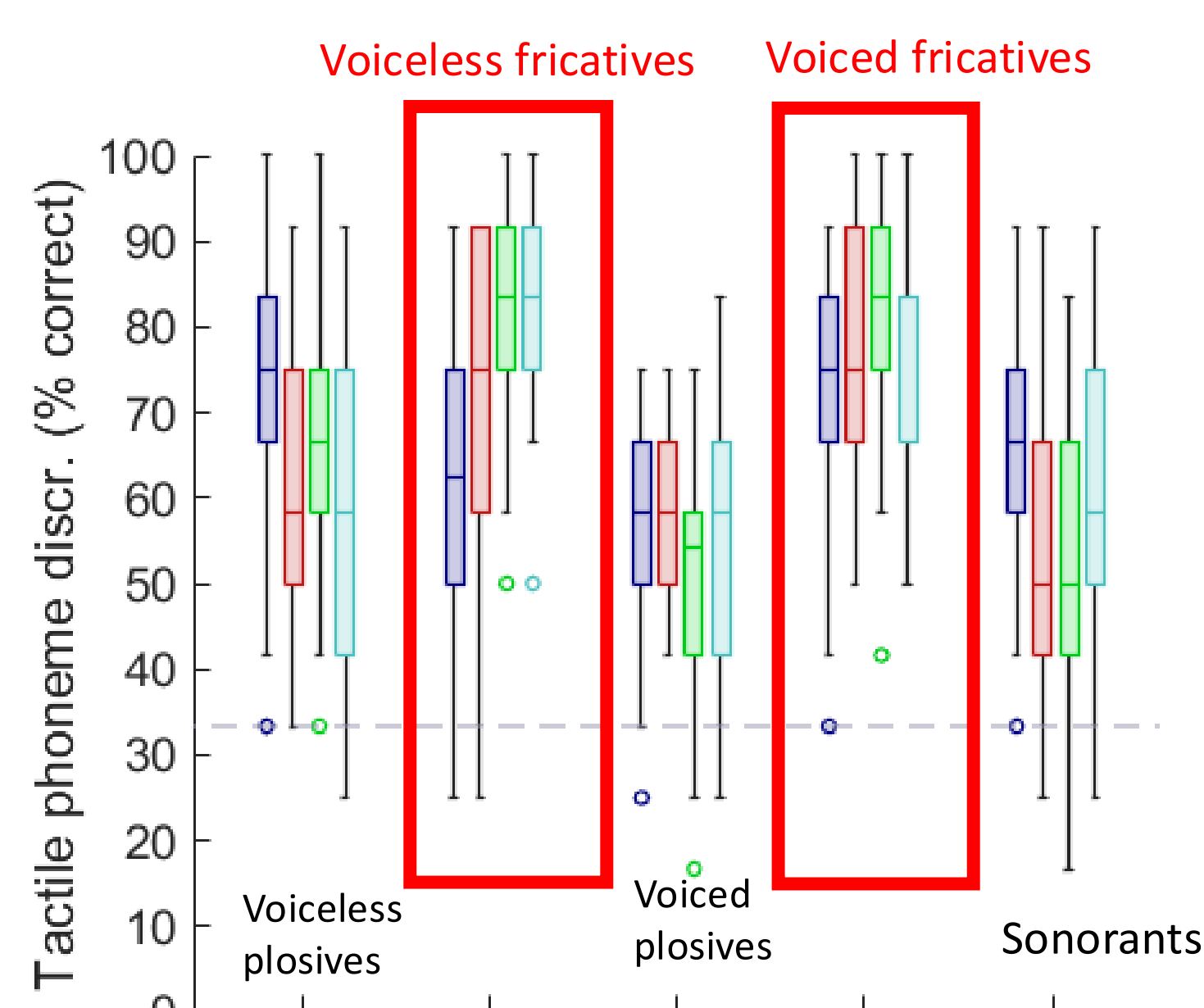


Figure 5. Box-and-whisker plot showing %correct for all algorithms across 5 consonant contrasts. Chance level at 33.3%

Results

Overall accuracy:

Algorithm	Mean %correct
Formant Focusing	71.56% (SD: 8.03%)
Tri-Focusing	63.87% (SD: 8.18%)
Mid-Focusing	66.09% (SD: 9.23%)
Mid-High Focusing	68.23% (SD: 9.40%)

Formant Focusing

- **Highest overall** and **vowel** %correct
- **Formant Focusing > Tri-Focusing**
- Mean difference:
 - **8.09%** ($p < .001$) (All phonemes)
 - **18.94%** ($p < .001$) (Vowels)

Tri-Focusing

- **Lowest overall** and **vowel** %correct
- **Highest fricative** %correct
- **Tri-Focusing > Formant Focusing**
- Mean difference:
 - **15.56%** ($p < .001$) (All fricatives)
 - **21.67%** ($p < .001$) (Voiceless fricatives)
 - **9.44%** ($p = .004$) (Voiced fricatives)
- Enhancing fricative discrimination with a trade-off from vowel discrimination

Mid-High Focusing & Mid-Focusing:

- No unique benefits identified

Discussion

1. Potential two-mode EHS strategy for optimisation:

- Formant Focusing = Default setting (✓ overall phonemes & ✓ vowels)
- Tri-Focusing = targeted benefits for ✓ fricatives
- Potential adaptive algorithms:
 - Real-time switching between algorithms, determined by incoming speech signals
- Technologically & theoretically possible

2. Partially fulfilled a research gap: Effective algorithm for ✓ fricatives & ✗ plosives

3. A phoneme-specific algorithm designed for targeted phoneme groups is feasible

4. Potential clinical usage = additional haptic cues for:

- **A) Poorer CI listeners**
- **B) Individuals without access to CI or HA**
- **C) Individuals with high-frequency sloping SNHL**
 - Possible difficulties of identifying fricatives even in quiet if severe
 - Significant concentration on fricative perception in noise

5. Further study: effective algorithms for plosives, EHS clinical trials on actual patients

Conclusion

- **Formant Focusing**: best for overall phonemes & vowels
- **Tri-Focusing**: best for fricatives
- Can potentially complement each other in a **two-mode EHS strategy**
- Mid-High Focusing & Mid-Focusing: no unique benefits
- Partially identified effective algorithms for obstruents (✓ fricatives & ✗ plosives)
- Targeted **phoneme-specific conversion algorithms** are **feasible**

References:

¹Fletcher, M.D. et al. (2024) 'Improved tactile speech perception using audio-to-tactile sensory substitution https://research.soton.ac.uk/ehs/ with formant frequency focusing', *Scientific Reports*, 14(1). Available at: https://doi.org/10.1038/s41598-024-55429-3.

