Analysis of complex sounds at high frequencies

Daniel R. Guest and Andrew J. Oxenham
Auditory Perception and Cognition Lab, Department of Psychology, University of Minnesota

Introduction

- Level discrimination and AM detection vary little as a function of frequency [1, 2]

- However, both of these tasks are simple and can be performed on the basis of cues from a single frequency channel / auditory-nerve fiber

- Whiteford et al. (2020) [3] recently showed that detecting incoherence in the modulator phases of two SAM tones worsened at high carrier frequencies

- Are similar deficits seen at high frequencies in other tasks thought to depend on cross-frequency comparisons?

Overview

Methods

- Measured psychophysical performance at low frequencies and high frequencies in multiple tasks

- Some tasks were designed to be possible using information only from a single channel (level discrimination, ripple detection)

- Other tasks were designed to require information from multiple frequency channels (profile analysis, ripple direction discrimination)

- We then related psychophysical performance to simulations of auditory-nerve responses [3]

Key questions

- Q1: Can listeners perform profile analysis at high frequencies?

- Q2: Can listeners perform ripple direction discrimination at high frequencies?

- Q3: Are patterns of psychophysical performance related to auditory-nerve coding?

Stimuli

Log-spaced complex tones

- Random-phase log-spaced complex tones

- Frequencies spaced from 0.6-1.6 kHz (low freq) or 6 to 16 kHz (high freq)

- Variable number of components (3, 5, 9, or 15)

- Either ...
  - fixed pedestal level of 60 dB SPL (level discrimination)
  - random pedestal level over 50-70 dB SPL (profile analysis)

- 350 ms in duration

Spectrotemporal ripples

- Sum of 300 random-phase SAM tones

- Ripple rate of 2 Hz

- Ripple density of 4 cycles/ octave

- Log-spaced carriers from 0.5 18 kHz at 45 dB SPL per-component

- Bandpass filtered from 0.6-1.6 kHz (low freq) or 6 to 17 kHz (high freq)

- 1000 ms in duration

Behavior

Deficits in profile analysis at high frequencies

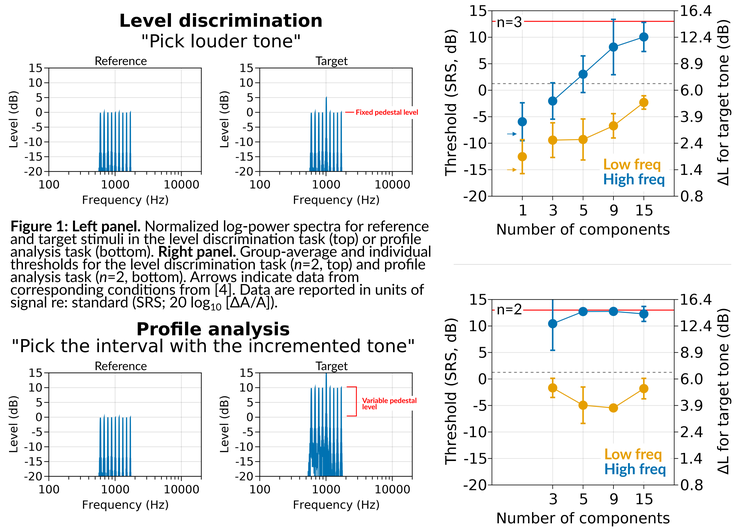


Figure 1: Left panel. Normalized log-power spectra for reference and target stimuli in the level discrimination task (top) or profile analysis task (bottom). **Right panel.** Group-average and individual thresholds for the level discrimination task (n=2, top) and profile analysis task (n=2, bottom). Arrows indicate data from corresponding conditions from [4]. Data are reported in units of signal re: standard (SRS; 20 $\log_{10}$ [ΔA/A]).

No deficit in ripple detection/discrimination at high frequencies

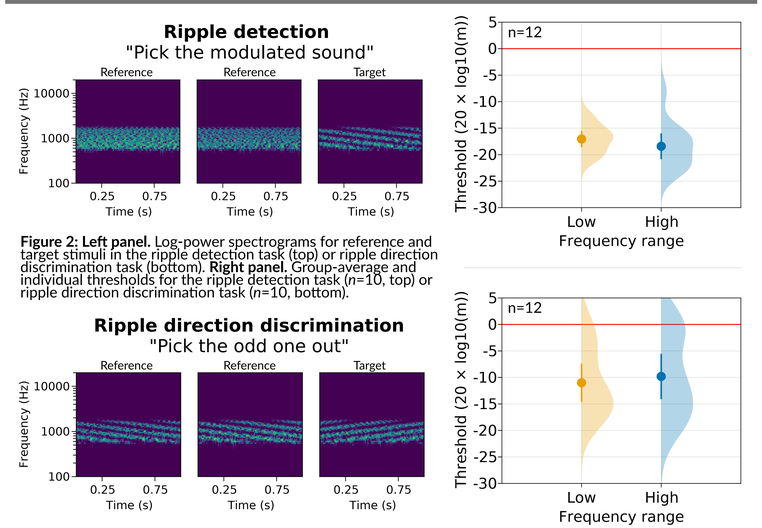
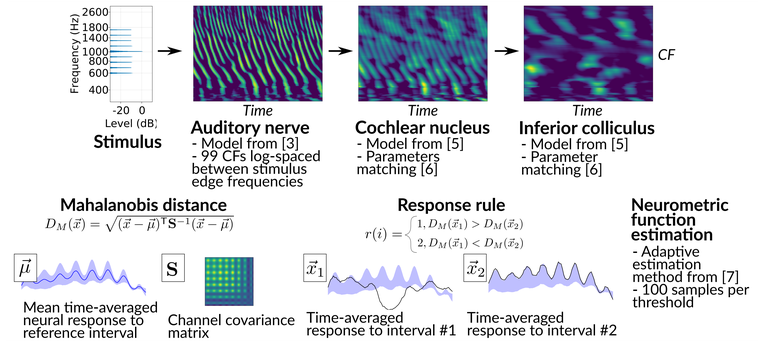


Figure 2: Left panel. Log-power spectrograms for reference and target stimuli in the ripple detection task (top) or ripple direction discrimination task (bottom). **Right panel.** Group-average and individual thresholds for the ripple detection task (n=10, top) or ripple direction discrimination task (n=10, bottom).

Modeling

Template-based models used to estimate thresholds



Neither AN nor midbrain rate decoding captured all trends

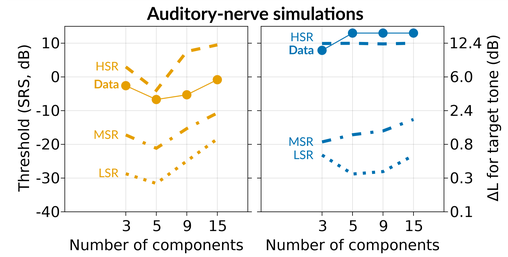


Figure 3: Threshold predictions from the Mahalanobis-distance template model based on decoding AN average firing rates. Data are indicated with points and model predictions with dashed lines. Different fiber types, as indicated by line style, yielded differing predictions.

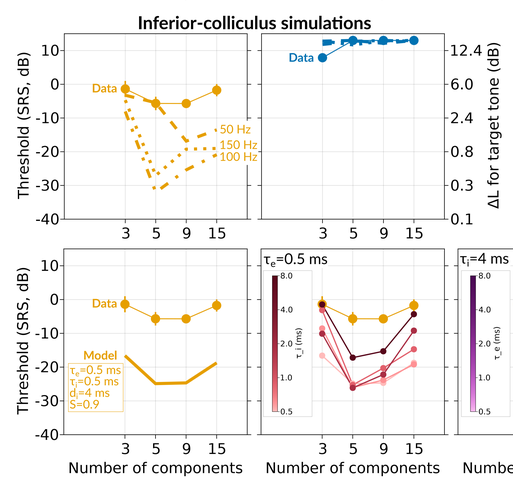


Figure 4: Thresholds as above, but based on decoding firing rates in the inferior colliculus. In the panels to the left, different dashed lines show predictions corresponding to parameter sets used in [5] to achieve nominal BMFs of 50, 100, and 150 Hz.

In the panels below, the inhibitory delay time was fixed at 4 ms and the excitatory and inhibitory time constants were swept over a range from 0.5 to 8 ms. Results are shown for a subset of resulting combinations.

Conclusions

- A1: Listeners could not perform profile analysis at high frequencies (Figure 1)

- A2: Listeners could perform ripple direction discrimination at high frequencies (Figure 2)

- A3: Template-based decoding of AN rates did not match behavioral trends in profile analysis (Figure 3)

- In AN simulations, MSR and LSR fibers outperformed listeners; HSR fiber performance was much poorer

- Shorter IC time constants needed to match data relative to similar simulations in [6]

- Future modeling will also explore representations of ripple stimuli in AN / IC

Acknowledgments

Funding: Work supported by UMN College of Liberal Arts Graduate Fellowship awarded to D.R.G., NIH R01 DC005216 awarded to A.J.O., NIH F31 DC019247-01 awarded to D.R.G., and NSF NRT-UtB1734815

Open source code/software:
- AFC [8]
- Julia (Parameters, Chain, Makie, DataFrames, Algebra of Graphics, DrWatson)
- Inkscape

References

[1] Florentine, M., Buus, S. R., & Mason, C. R. (1987). Level discrimination as a function of level for tones from 0.25 to 16 kHz. *The Journal of the Acoustical Society of America*, *81*(5), 1528-1541.

[2] Sek, A., & Moore, B. C. (1995). Frequency discrimination as a function of frequency, measured in several ways. *The Journal of the Acoustical Society of America*, *97*(4), 2479-2486.

[3] Whiteford, K. L., Kreft, H. A., & Oxenham, A. J. (2020). The role of cochlear place coding in the perception of frequency modulation. *Elife*, *9*, e58468.

[4] Zilany, M. S., Bruce, I. C., & Carney, L. H. (2014). Updated parameters and expanded simulation options for a model of the auditory periphery. *The Journal of the Acoustical Society of America*, *135*(1), 283-286.

[5] Nelson, P. C., & Carney, L. H. (2004). A phenomenological model of peripheral and central neural responses to amplitude-modulated tones. *The Journal of the Acoustical Society of America*, *116*(4), 2173-2186.

[6] Maxwell, B. N., Richards, V. M., & Carney, L. H. (2020). Neural fluctuation cues for simultaneous notched-noise masking and profile-analysis tasks: Insights from model midbrain responses. *The Journal of the Acoustical Society of America*, *147*(5), 3523-3537.

[7] Kontsevich, L. L., & Tyler, C. W. (1999). Bayesian adaptive estimation of psychometric slope and threshold. *Vision Research*, *39*(16), 2729-2737.

[8] Ewert (2013), AFC - A modular framework for running psychoacoustic experiments and computational perception models. *Proceedings of the International Conference on Acoustics*. 1326-1329.

Supporting materials

Poster available at:

https://guestdaniel.github.io/download/GuestOxenhamARO2022.pdf