Pitch perception of concurrent high-frequency complex tones

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

Introduction

- Pitch perception of resolved complex tones can remain fairly accurate even when all harmonics are beyond putative limits of phase locking [9, 4, 2]
- Pitch perception of complex tones can also remain fairly accurate in the presence of complex tone maskers [6, 5, 10]
- However, is is unknown whether accurate pitch perception is possible with both (1) complex tone maskers and (2) targets entirely beyond the limits of phase locking

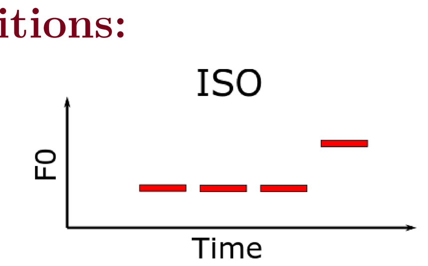
Summary

- **Paradigm:** Listeners heard three tones with same F0 (reference) followed by one tone with different F0 (target) mixed with maskers and indicated direction of F0 change
- **Experiments:**
  - **Exp. 1a and Exp. 1b:** F0DLs w/ and w/o masker tone
  - **Exp. 2:** Percent correct at fixed interval w/ two masker tones
  - **Exp. 3:** Target-to-masker ratio (TMR) required for fixed interval w/ two masker tones

Methods

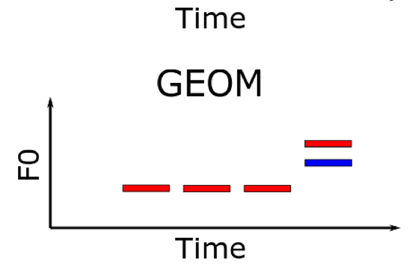
- **Targets:** Complex tones in TEN noise [8]
  - **Exp. 1a and Exp. 2:** harmonics 6-10 of F0
  - **Exp. 1b and Exp. 3:** all harmonics of F0, bandpass filtered (12th order Butterworth, cutoffs at 5.5× and 10.5× nominal F0)
  - 50 ± 3 dB SPL rove per component (pre-filtering)
- **Maskers:** Complex tones
  - **Exp. 1a and Exp. 2:** harmonics 5-11 of F0
  - **Exp. 1b and Exp. 3:** all harmonics of F0, bandpass filtered (12th order Butterworth, cutoffs at 4× and 12× nominal F0)
  - 50 ± 3 dB SPL rove per component (pre-filtering)
- **Frequency conditions:**
  - Low Freq (nominal F0 = 280 Hz ± 10% rove)
  - High Freq (nominal F0 = 1400 Hz ± 10% rove)
- **Masker conditions:**

Exp. 1a & Exp. 1b



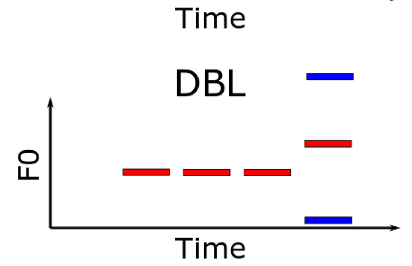
No masker tone

Exp. 1a & Exp. 1b



Masker tone geometrically centered between target and reference

Exp. 2 & Exp. 3



One masker tone above and one masker tone below target (at least 5.25 semitones)

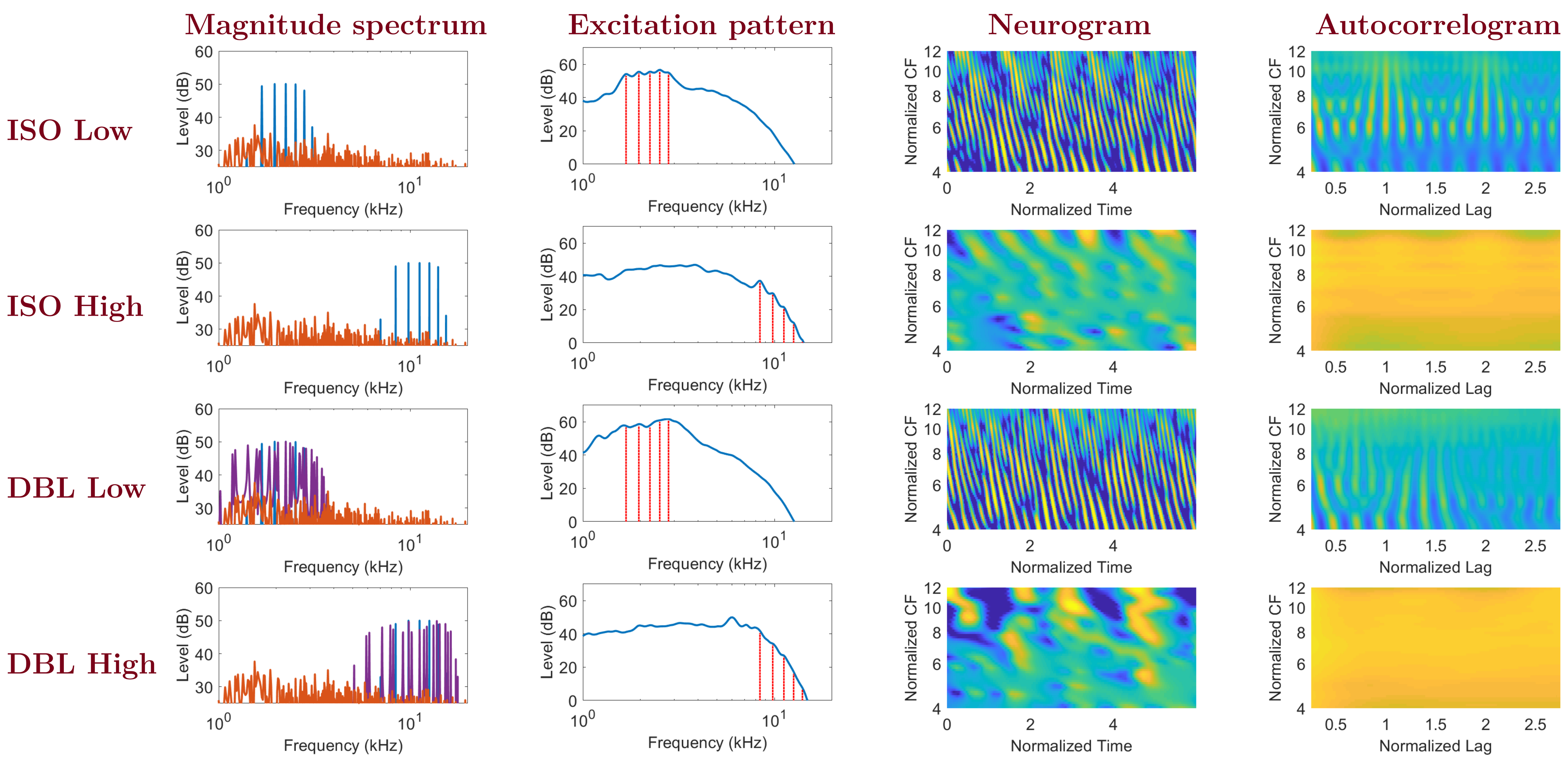
- **Interval sizes:**
  - **Exp. 2:** 2× F0DL from Exp. 1a ISO (within [1.5, 2.5] semitones)
  - **Exp. 3:** 1.5× and 2.5× F0DL from Exp. 1b ISO

Hypotheses

- **H1:** Pitch perception will be poorer (although still good) in High Freq than Low Freq [4]
- **H2:** Single masker will have a larger detrimental impact on performance in High Freq than Low Freq
- **H3:** Higher TMRs needed for the same performance in High Freq than Low Freq

Stimuli

Figure 1: Magnitude spectra, excitation patterns, neurograms, and neural autocorrelograms of the ISO and DBL stimuli. Details regarding the simulations are available in the right-hand "Computational Modeling" column.



Results — Exp. 1 (ISO and GEOM)

- High Freq worse than Low Freq
- GEOM masker had larger effect on Low Freq than High Freq

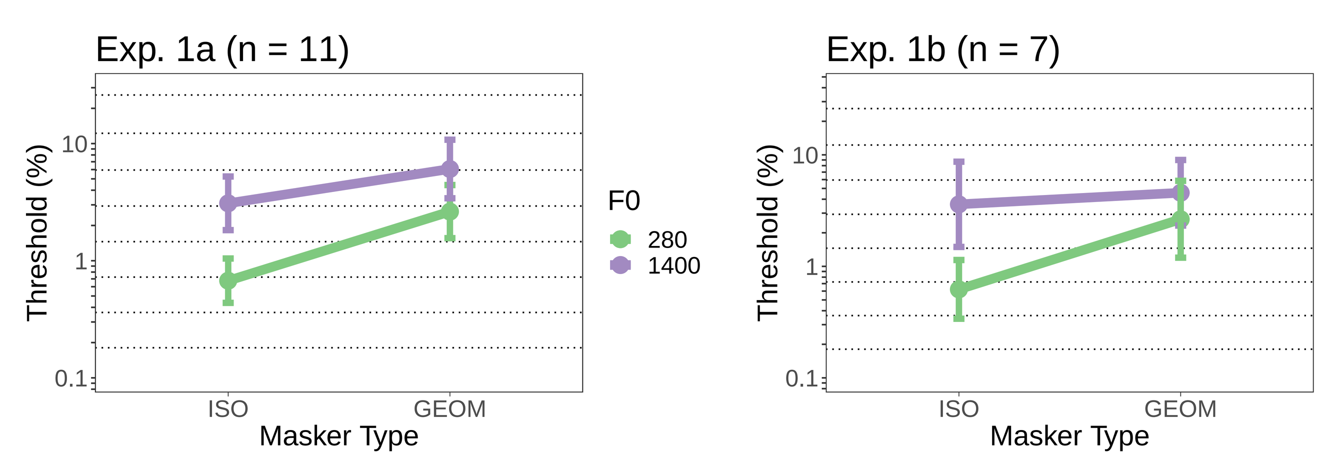


Figure 2: Results from 7 runs per condition for Exp. 1a and Exp. 1b. Error bars indicate ±1.96 SEM.

Results — Exp. 2 (DBL)

- Above chance performance for both conditions
- High Freq worse than Low Freq

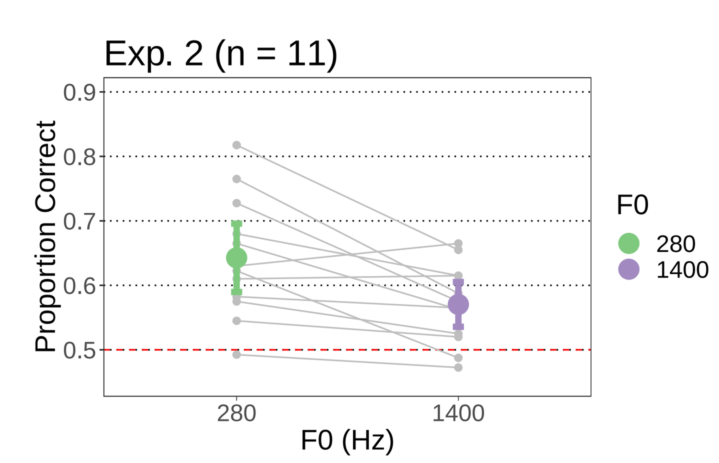


Figure 3: Results from 400 trials per condition for Exp. 2. Error bars indicate ±1.96 SEM. Masker type for all conditions was DBL.

Results — Exp. 3 (DBL)

- Higher TMRs appear to be required in High Freq than Low Freq for same performance

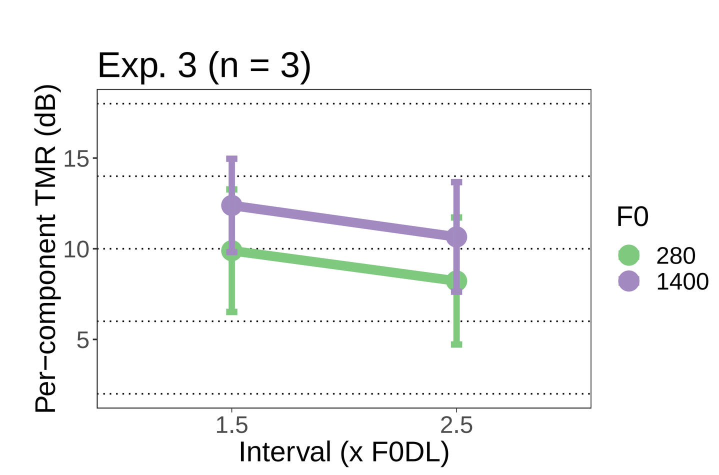


Figure 4: Results from 7 runs per condition for Exp. 3. Error bars indicate ±1.96 SEM. Masker type for all conditions was DBL.

AN Ideal Observer (ISO)

- **All-information predictions overestimate decrease from Low Freq to High Freq**
- **Rate-place predictions underestimate decrease from Low Freq to High Freq**

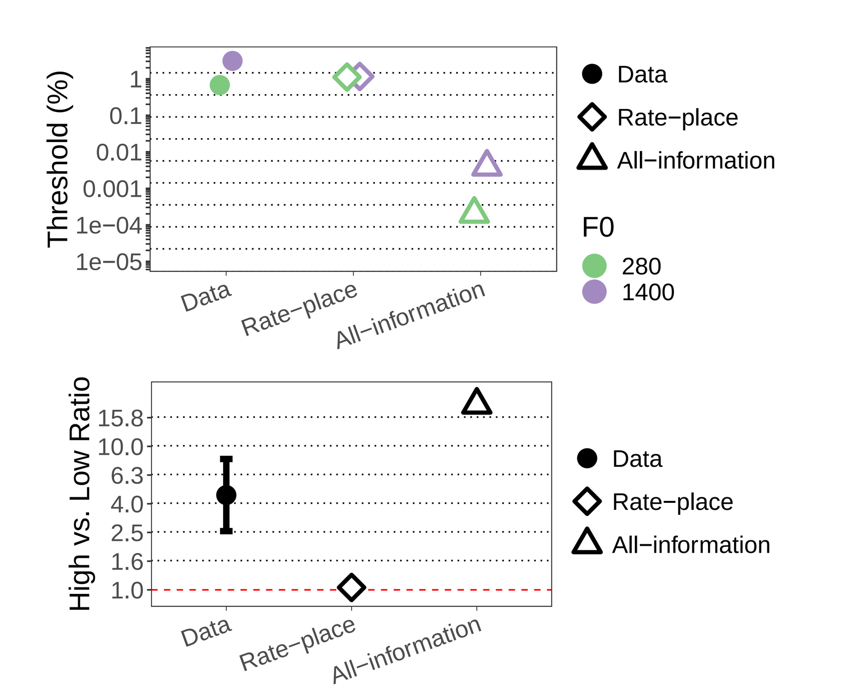


Figure 5: Predicted thresholds for an ideal observer based on simulated auditory nerve activity for the ISO stimulus. Comparison behavioral data is from Exp. 1a. Error bars indicate ±1.96 SEM.

Computational modeling

- Excitation patterns in Figure 1:
  - Output of 256 auditory filters described in Glasberg and Moore [1] as a function of characteristic frequency (CF) from 0.20 to 20 kHz
  - Role of outer and middle ear included according to Moore and Glasberg [7]
- Neurograms and autocorrelograms in Figure 1:
  - Firing rates of auditory nerve (AN) model of Zilany, Bruce, and Carney [11] as a function of CF/F0 and time/(1/F0)
  - 256 characteristic frequencies from 0.20 to 20 kHz
  - Low spontaneous rate fibers with Glasberg and Moore tuning
- Predicted thresholds in AN Ideal Observer:
  - AN model and ideal observer of Heinz, Colburn, and Carney [3]
    - $\mathrm{JND}_{\text{All-information}} = \left(\sum_i \int_0^T \frac{1}{r_i(t,f)} \left[\frac{\partial}{\partial f} r_i(t,f) \right]^2 dt \right)^{-1/2}$
    - $\mathrm{JND}_{\text{Rate-place}} = \left(\sum_i \int_0^T \frac{1}{\bar{r}_i(f)} \left[\frac{\partial}{\partial f} \bar{r}_i(f) \right]^2 dt \right)^{-1/2}$
    - $r_i(t,f)$ is the instantaneous firing rate of fiber i at time t with stimulus F0 f
  - 120 characteristic frequencies from 0.20 to 20 kHz
    - Only responses between 5× and 12× nominal F0 were used (to simulate the effect of TEN noise limiting audibility outside bandpass region)
  - 100 high spontaneous rate fibers per characteristic frequency

Conclusions

- **H1:** ✓ Pitch perception was poorer in High Freq than Low Freq, but performance both with and without maskers was still good (i.e., F0DLs < 1 semitone)
- **H2:** ✗ Single masker unexpectedly had larger detrimental impact in Low Freq than High Freq
- **H3:** ✓ Larger TMRs appear to be required for High Freq than Low Freq for same performance
- **Modeling:** Neither rate-place nor all-information predictions match decrease from Low Freq to High Freq in behavioral data

Significance

- Accurate pitch perception of mixtures of complex tones at high frequencies is possible
- Neural mechanisms of high-frequency complex pitch perception remain unclear
  - Simulations with more accurate models of the auditory periphery may provide further insight

Acknowledgements

- This research was funded by a UMN College of Liberal Arts Graduate Fellowship awarded to D.R.G., NIH R01 DC005216 awarded to A.J.O., and NSF NRT-UtB1734815

Bibliography

[1] Glasberg, B. R. and Moore, B. C. J. In: *Hearing Research* 47 (1990), pp. 103–138. DOI: 10.1016/0378-5955(90)90170-T.

[2] Gockel, H. E. and Carlyon, R. P. In: *Acta Acustica* 104.5 (2018), pp. 766–769. DOI: 10.3813/AAA.919219.

[3] Heinz, M. G., Colburn, H. S., and Carney, L. H. In: *Neural Computation* 13 (2001), pp. 2273–2316. DOI: 10.1162/089976601750541813.

[4] Lau, B. K., Mehta, A. H., and Oxenham, A. J. In: *The Journal of Neuroscience* (2017). DOI: 10.1523/JNEUROSCI.1507-17.2017.

[5] Micheyl, C., Bernstein, J. G. W., and Oxenham, A. J. In: *The Journal of the Acoustical Society of America* 120.3 (2006), pp. 1493–1505.

[6] Micheyl, C., Keebler, M. V., and Oxenham, A. J. In: *The Journal of the Acoustical Society of America* 128.1 (2010), pp. 257–269. DOI: 10.1121/1.3372751.

[7] Moore, B. C. J. and Glasberg, B. R. In: *Hearing Research* 28 (1987), pp. 209–225.

[8] Moore, B. C. J. et al. In: *British Journal of Audiology* 34.4 (2000), pp. 205–224. DOI: 10.3109/03005364000000131.

[9] Oxenham, A. J. et al. In: *Proceedings of the National Academy of Science* 108.18 (2011), pp. 7629–7634. DOI: 10.1073/pnas.1015291108.

[10] Wang, J. et al. In: *The Journal of the Acoustical Society of America* 132.1 (2012), pp. 339–356. DOI: 10.1121/1.4728165.

[11] Zilany, M. S. A., Bruce, I. C., and Carney, L. H. In: *The Journal of the Acoustical Society of America* 135.1 (2014), pp. 283–286. DOI: 10.1121/1.4837815.