Improving the Detection of Melodic Sequences Through the Addition of Inharmonic Frequencies

Connor Wessel *[1], **Michael Schutz** †[1,2]
[1]Department of Psychology, Neuroscience, and Behaviour, McMaster University, Hamilton, Ontario,Canada
[2]School of the Arts, McMaster University, Hamilton, Ontario, Canada

1 Introduction

Our auditory system is continually challenged to extract signals informing us about our environment, filtering out noise and irrelevant signals. Detection and discrimination are two processes that are crucial in this endeavor. The former is our ability to perceive specific sounds, whereas the latter is distinguishing between signals.

Detection can be influenced by perceptual grouping of sounds in our environment. Duplex perception is an example recorded in linguistics literature, in which a 'chirp' presented binaurally alongside a verbal consonant can be simultaneously perceived as contributing to and independent from the consonant [1]. This can also apply to non-verbal sounds and to harmonic tones with a single-mistuned harmonic [2, 3].

Since duplex perception can trigger through mistuning harmonics, this creates a gateway through which to explore its role in detectability. Inharmonicity can be beneficial for grabbing attention, thereby leading to greater detection [4, 5]. Consequently, introducing inharmonic frequencies into otherwise simple harmonic tones could aid detection of these simple tones through the partial perceptual binding established by duplex perception. This study explored the possibility by measuring participants' ability to identify melodic sequences with additional, inharmonic frequencies in noise.

2 Method

2.1 Participants

We recruited sixty students at McMaster University, who received course credit as compensation for their time. Participants completed the study online and received instruction to use a laptop/desktop computer with headphones.

2.2 Materials

We designed three tones for this experiment, each of which contains two components: the harmonically simple 'base' tone, and the complex higher harmonics. The base tones contain three harmonics, including the fundamental, and two overtones along the harmonic series. The higher pitched harmonics contain 10 frequencies, starting at 2000 Hz, and increasing at a rate of 1000 Hz – n x 100 Hz. Each tone lasts 500 ms, starting with a 5 ms linear ramp and ending with a 495 ms exponential decay.

We manipulated three aspects of the sounds, including i) the melodic contour of the higher-pitched harmonics, ii) the direction of the sequence, and iii) the signal-to-noise ratio (SNR). For the first, the tones could either be 'tracking' (follow the same sequence as the base tones), 'stationary' (each tone is the same pitch), or 'absent' (higher harmonics omitted). For the second, we arranged the tones into a three-tone sequence that either ascended or descended. Tracking higher harmonics also ascended and descended at the same frequency interval ratio. Lastly, we presented sequences at six different SNRs, ranging from -20 dB to -30 dB. Figure 1 provides an example of an ascending sequence.

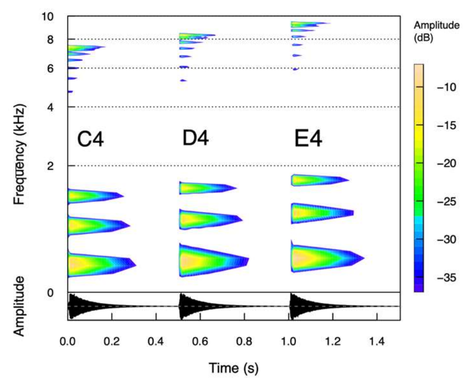

Figure 1: Spectrogram of the three tones within the ascending sequence. The horizontal axis displays time (s), and the vertical axis displays frequency (kHz). The three frequencies at the bottom between 0-2 kHz represent the 'base' sequence, while the frequencies between 4-10 kHz represent the 'higher harmonics.' The labels represent the musical pitch of each tone.

2.3 Procedure

The dependent variables for this experiment include the proportion of correctly identified sequences (ascending or descending) and average reaction times in each trial (ms). This experiment followed a 3 (higher harmonic type) x 2 (sequence) within-subjects design.

We started by presented participants with examples of the stimuli. Participants received instructions to guess the sequence of just the base tones and not the higher harmonics. After completing practice trials, participants underwent 180 experimental trials. This involved presenting the target sound simultaneously with background noise. For 3000 ms afterward, participants received instruction to press 'a' if they heard an ascending sequence, or 'd' if they heard a descending sequence. Failure to respond marked a trial as a miss, leading to its exclusion from analyses.

* wesselc@mcmaster.ca
† schutz@mcmaster.ca

3 Results

3.1 Accuracy

We define accuracy as the average proportion of correctly identified sequences within a given condition. We analysed this data using a two-way within-subjects ANOVA, using Greenhouse-Geiser corrections where required. Figure 2 presents the proportion of correct trials split across the presence of higher harmonics.

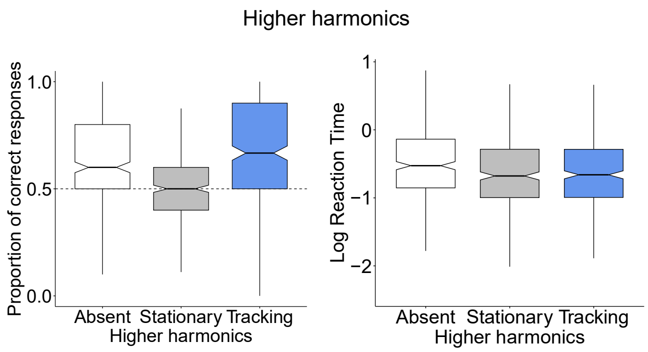

Figure 2: Boxplot displaying the average proportion of correct responses (left) and log reaction time (right) for each type of higher harmonics. The dotted line for proportion represents chance performance (0.5).

We found the presence of higher harmonics significantly affected accuracy, $GG(114) = 0.84$, $p < .001$. Tracking harmonics produced the highest accuracy (M = 0.68, SD = 0.47) and stationary produced the lowest (M = 0.49, SD = 0.50), while their absence falls in the centre (M = .62, SD = .50). A subsequent Tukey test found that comparisons between all types yielded significant differences.

3.2 Response time

We took the average, log-transformed time to press a key in each trial. While we used log transformed values in the analysis, we report raw values for clarity. Figure 2 presents the average response times across the presence of higher harmonics.

We found a significant effect of the presence of higher harmonics, $F(114) = 19.58$, $p < .001$, with lower response times for stationary (M = 566 ms, SD = 528 ms) and tracking (M = 568 ms, SD = 519 ms) vs. absent (M = 667 ms, SD = 546 ms) harmonics. A subsequent Tukey test found that absent sequences yielded significantly different response times from tracking and stationary but tracking and stationary did not differ from each other.

4 Discussion

Adding higher harmonics to a melodic sequence reduced response time and led to context specific effects on accuracy. Specifically, accuracy increased when higher harmonics tracked the melodic sequence of the base tones and decreased when they remained stationary.

Overall, the tracking higher harmonics improve detectability while stationary harmonics reduce it to chance level. This implies that the higher harmonics overwhelm perception of the base tones. Presumably, the tracking higher harmonics lead to better performance as they contain the task-relevant information, whereas the stationary harmonics block it out. This could also be a matter of attention being drawn to the higher harmonics and not to the base tones.

In terms of duplex perception, it is unclear whether the partial perception binding occurred due to the potential overshadowing. This could be addressed by either (1) adjusting the intensity of the higher harmonics relative to the base tones or (2) introducing greater separation between the base tones and higher harmonics by manipulating further parameters of the sound, such as amplitude envelope. If differences in accuracy emerge between different degrees of separation, this can suggest the extent to which higher harmonics bind with the base tones. In terms of duplex perception, we can test the limits on how varied two components of a sound can be to produce dual states of perception.

5 Conclusion

Adding higher harmonics to tones can provide context dependent boosts in detectability. This relies on the higher harmonics containing the information through which participants derive required meaning. While the question of how partial perceptual bindings plays a role will require further exploration, the boost in detectability for tracking higher harmonics is meaningful in practical contexts, such as alert design.

Acknowledgments

This work was supported by Natural Sciences and Engineering Research Council of Canada (NSERC), the Canadian Foundation for Innovation, and the Auditory Performance Division of the US Navy and the Office of Naval Research Grant N00014-22-1-2184.

References

[1] Liberman AM, Isenberg D, Rakerd B. Duplex perception of cues for stop consonants: Evidence for a phonetic mode. *Perception & Psychophysics* 1981; **30**(2): 133–143

[2] Fowler CA, Rosenblum LD. Duplex perception: A compar-ison of monosyllables and slamming doors. Journal of Ex-perimental Psychology: Human Perception and Perfor-mance 1990; 16(4): 742–754

[3] Moore BCJ, Peters RW, Glasberg BR. Thresholds for the detection of inharmonicity in complex tones. The Journal of the Acoustical Society of America 1985; 77(5), 1861–1867

[4] Bonin T, Smilek D. Inharmonic music elicits more negative affect and interferes more with a concurrent cognitive task than does harmonic music. Atten Percept Psychophys 2016; 78(3): 946–959

[5] McPherson MJ, Grace RC, McDermott JH. Harmonicity aids hearing in noise. Atten Percept Psychophys 2022; 84(3): 1016–1042.