ÖAW AUSTRIAN ACADEMY OF SCIENCES

Perception of simultaneous versus sequential musical harmony in cochlear implant listeners

Marie-Luise Augsten, Martin Lindenbeck, Bernhard Laback

ARI

Experiment 1: Harmony perception with real-world musical stimuli

Motivation

Musical harmony perception in cochlear implant (CI) listeners is severely impaired. Mostly, musical harmonies consist of a superposition of multiple simultaneous voices [1]. These broadband sounds lead to complex interactions between CI electrodes and are thus difficult to perceive by CI listeners (e.g., [2]). One technique for unraveling complexity might be a sequential rather than simultaneous presentation of chord voices (e.g., arpeggio-like; Fig. 1). We investigated harmony perception of simultaneously and sequentially harmonized musical real-world sequences in which single chords were harmonically altered and hypothesized that *discrimination of altered vs. non-altered harmonies is better for sequentially harmonized sequences.*

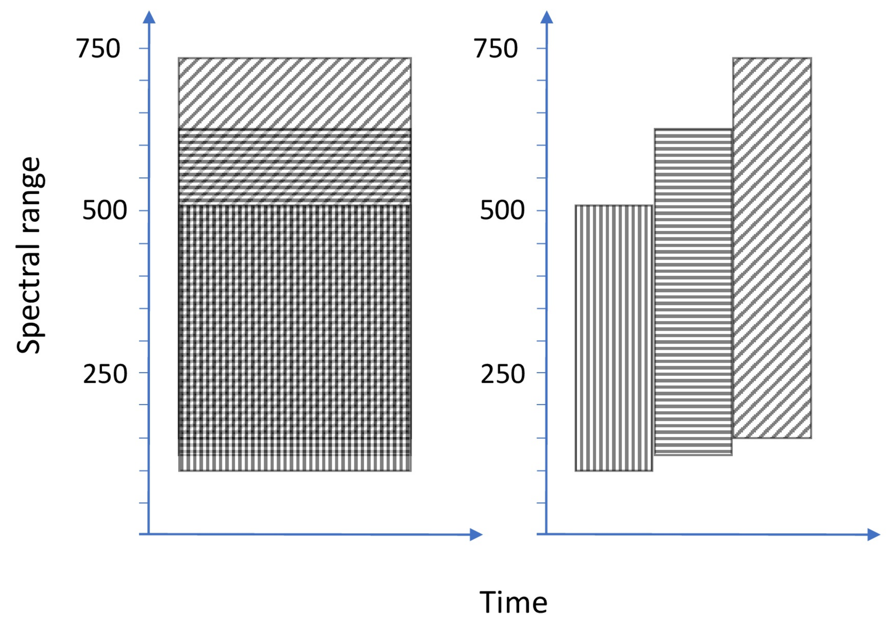

Fig. 1

Schematic illustration of a simultaneously (left) and sequentially (right) presented major triad. Time is depicted on the abscissa and spectral range on the ordinate. The triad is composed of five-component HC tones with F0s of 100 Hz (root), 126 Hz (major third), and 150 Hz (fifth). For the simultaneously presented triad spectral ranges of HC tones largely overlap whereas the overlap is avoided for the sequentially presented triad.

Musical excerpts of the song "Der Mond ist aufgegangen", in simultaneous (**A**) and sequential (**B**) harmonization. The last chord is altered harmonically either to an in-scale (i.e., consonant) chord (top, e.g., a minor chord) or to an out-of-scale (i.e., dissonant) chord (bottom). Both alterations are unexpected, however, the out-of-scale chord is perceived as dissonant and thus can be processed on a sensory level whereas the in-scale chord is perceived as consonant and can be processed on a syntactical (i.e., structural) level only.

Fig. 2

Method

- Short real-word musical excerpts composed of five-component (including F0) harmonic complex (HC) tones (cf. [2-3])
- Simultaneous vs. sequential harmonization (Fig. 2)
- Harmonic alterations at various places: sensory (out-of-scale/dissonant) vs. syntactic (in-scale/consonant; Fig. 2)
- Participants assess 'correctness' of sequences (yes/no) and receive feedback
- Two MED-EL CI listeners with FS4 processing strategy [4]
- ≥ 80 repetitions per condition, sensitivity measure is d' [5]

Results

- CI116: Some sensitivity ($d' < 1$) for sequential out-of-scale alterations in closed position ($p < .05$)
- CI122: Some sensitivity ($d' < 1$) for simultaneous out-of-scale alterations in open position ($p < .01$)
- No sensitivity for in-scale alterations ($d' \sim 0$)

Conclusions

- Slight but significant sensitivity to out-of-scale alterations
- Sensory but not syntactical processing
- No advantage of sequential over simultaneous harmonization

Experiment 2: Interval and triad discrimination

Motivation

Experiment 1 showed no advantage for sequential presentation of harmony. This might be due to difficulties in integrating harmonic components which are dispersed over a relatively long time (cf. Fig. 2). Furthermore, the cognitive load elicited by such complex musical stimuli might have been too high to ensure sufficient salience of alterations (cf. [6]). Therefore, we investigated discrimination of single intervals and triads presented simultaneously and sequentially with shorter time intervals between the voices. We hypothesized that *discrimination of intervals and triads is better for sequential presentation and shorter duration.*

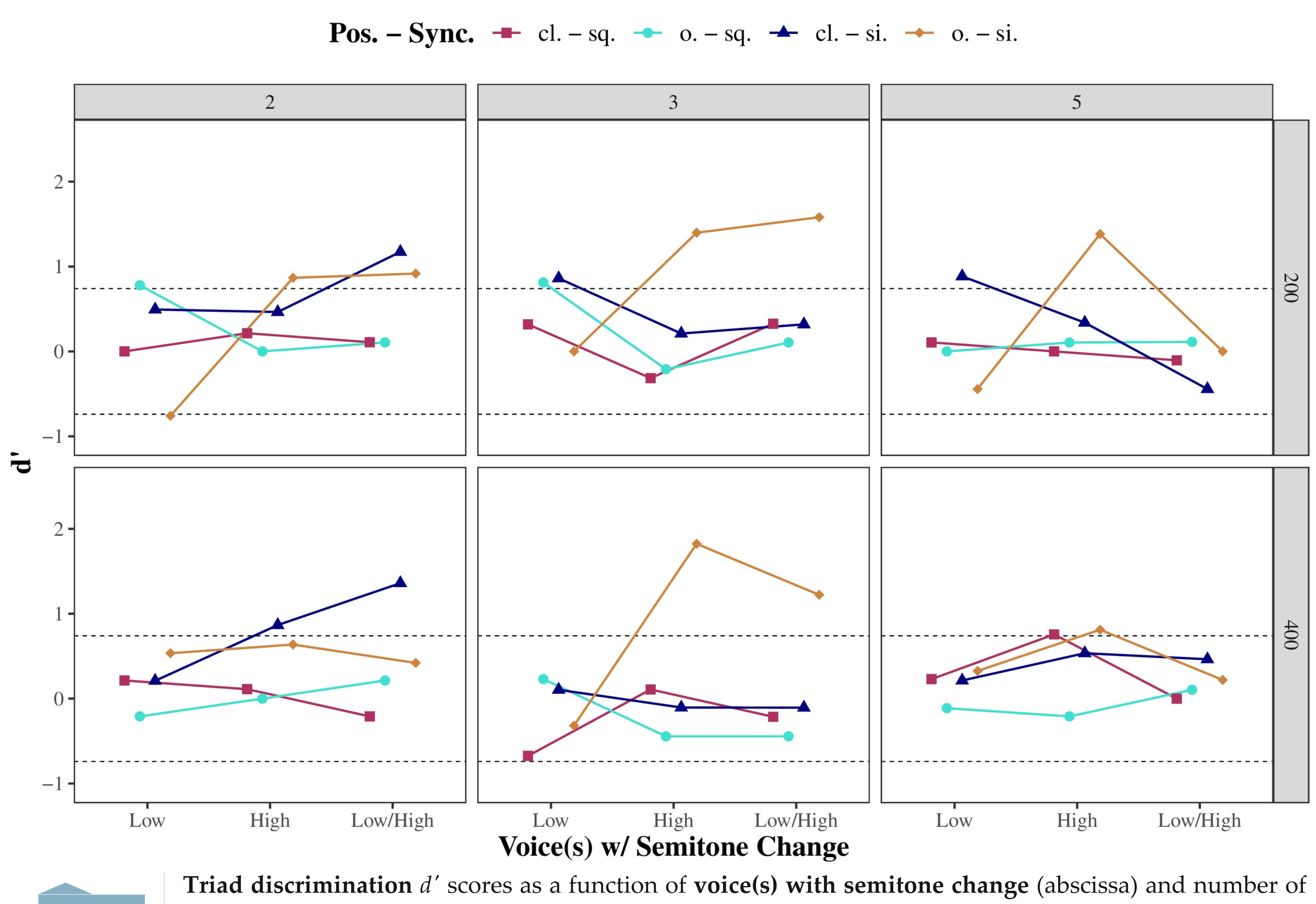

Fig. 3

Triad discrimination d' scores as a function of **voice(s) with semitone change** (abscissa) and number of harmonic complex **(HC) components** (columns), tested for **closed vs. open position** and in **simultaneous or sequential presentation** (colored lines). Two different total **stimulus durations** (rows, in ms) were tested with 500-ms gaps. Durations and gaps in sequential presentations matched simultaneous presentations, with no gap between voices within a sequential stimulus. For intervals, durations matched either to simultaneous or to sequential triads. For the sake of brevity, intervals are not depicted here. In total, F0s ranged from 117 to 311 Hz. 48 repetitions per condition were tested.

Conclusions

- Simultaneous presentation yields better performance than sequential presentation in all conditions
  – Role of beating between simultaneous voices?
- Open position of stimuli yields better performance
- Three-component stimuli lead to best results
- No detrimental effect of third voice (triads vs. intervals)
- No effect of duration

Method

- Intervals (two voices) and triads (three voices) composed of HC tones
- Varied stimulus parameters (cf. Fig. 3): Temporal synchrony, # of HC components (including F0), total stimulus duration
  - – For triads only: Position, Voice w/ change
  - – For interval-specific parameters see poster by Lindenbeck et al.
- 2I-2AFC task (same/different) with feedback
- One CI listener, rest as above

Results

Data were analyzed with binomial logistic regression [7-8].

Triad discrimination:

- Voice w/ change * Position * Synchrony * # Components ($p < .05$; Fig. 3)
  - – No advantage of sequential presentation
  - – Higher performance for simultaneous three-component triads with changes in only high or both high and low voice
  - – No effect of duration

Interval and triad discrimination (merged, matched for duration):

- Voice/F0 Range * Synchrony * Position ($p < .05$; Fig. 4)
  - – For both stimulus types (intervals and triads), performance is best for open position in simultaneous presentation and high F0 range/change in highest voice
- Stimulus Type * Position * Synchrony ($p < .05$; Fig. 5)
  - – Best performance for simultaneous stimuli in open position
  - – In open position, intervals yield largest synchrony difference

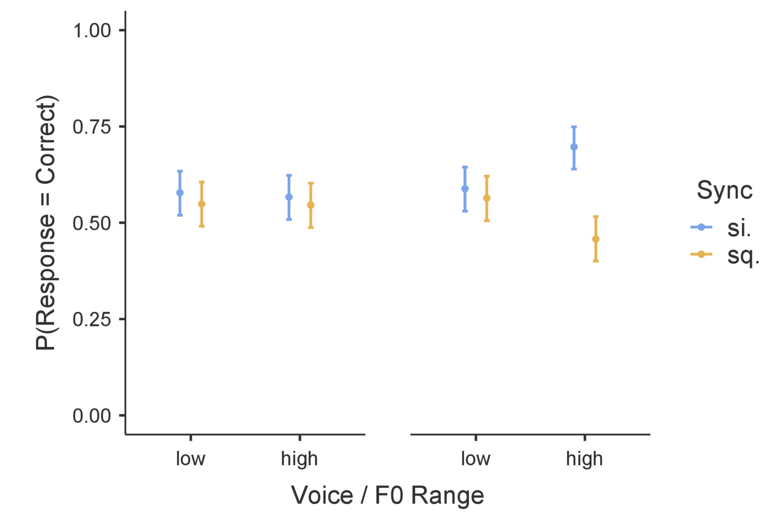

Fig. 4

Logistic regression results of interval and triad discrimination. The variables voice w/ change (for triads) and F0 range (for intervals) are merged. Also, the variables position (for triads) and reference interval (for intervals) are merged such that a high reference interval corresponds to the open position. The proportion of correct responses is depicted as a function of **voice w/ change/F0 range**, seperately for **closed vs. open position** (plots) and **simultaneous vs. sequential presentation** (colors).

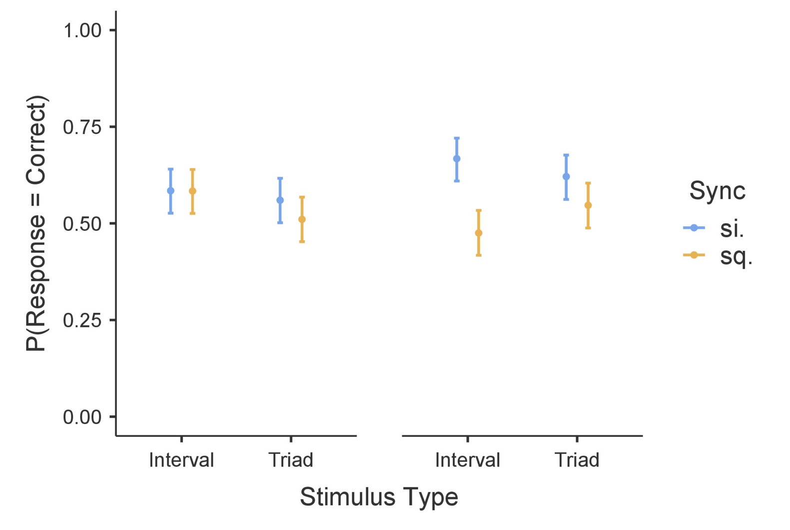

Fig. 5

Logistic regression results of interval and triad discrimination. Merging of variables as described in Fig. 4. The proportion of correct responses is depicted as a function of **stimulus type**, separately for **closed and open position** (plots) and **simultaneous vs. sequential presentation** (colors).

References

[1] Dahlhaus, C. (1980). Harmony. In S. Sadie (Ed.), The new Grove's dictionary of music and musicians (Vol. 8, pp. 175–177). Macmillan.
[2] Knobloch, M., Verhey, J. L., Ziese, M., Nitschmann, M., Arens, C., & Böckmann-Barthel, M. (2018). Musical Harmony in Electric Hearing. Music Perception, 36(1), 40–52. https://doi.org/10.1525/MP.2018.36.1.40.
[3] Zimmer, V., Verhey, J. L., Ziese, M., & Böckmann-Barthel, M. (2019). Harmony perception in prelingually deaf, juvenile cochlear implant users. Frontiers in Neuroscience, 13, 466. https://doi.org/10.3389/fnins.2019.00466
[4] Riss, D., Hamzavi, J.-S., Blineder, M., Honeder, C., Ehrenreich, I., Kaider, A., Baumgartner, W.-D., Gstoettner, W., & Arnoldner, C. (2014). FS4, FS4-p, and FSP: A 4-month crossover study of 3 fine structure sound-coding strategies. Ear and Hearing, 35(6), e272-e281. https://doi.org/10.1097/aud.0000000000000063
[5] Macmillan, N. A., & Creelman, C. D. (2005). Detection theory: a user's guide (2nd ed.). Lawrence Erlbaum Associates.
[6] Petersen, B., Andersen, A. S. F., Haumann, N. T., Højlund, A., Dietz, M. J., Michel, F., Riis, S. K., Brattico, E., & Vuust, P. (2020). The CI MuMuFe - A New MMN Paradigm for Measuring Music Discrimination in Electric Hearing. Frontiers in Neuroscience, 14, 2. https://doi.org/10.3389/fnins.2020.00002.
[7] Jamovi. (2022). Sydney, Australia: The jamovi project. Retrieved from https://jamovi.org
[8] R Core Team. (2021). R: A language and environment vor statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://cran.r-project.org

Supported by the Austrian Academy of Sciences with a DOC Fellowship (A-25606), by its Dr. Anton Oelzelt-Newin Fund, and by MED-EL.

ARI – ACOUSTICS RESEARCH INSTITUTE