DOI: 10.61782/fa.2023.0899

forum acusticum 2023

Music emotion perception with a cochlear implant: Can valence perception be predicted by other metrics?

Eleanor E Harding[1,2,3*] **Etienne Gaudrain**[1,4] **Robert Harris**[1,3]
Barbara Tillmann[4,5] **Bert Maat**[1,6] **Rolien Free**[1, 2,6] **Deniz Başkent**[1,2]
[1] Department of Otorhinolaryngology/Head and Neck Surgery, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands
[2]Graduate School of Medical Sciences, Research School of Behavioural and Cognitive Neurosciences, University of Groningen, Groningen, The Netherlands
[3]Prins Claus Conservatoire, Hanze University of Applied Sciences, Groningen, The Netherlands
[4]Lyon Neuroscience Research Center, CNRS UMR5292, Inserm U1028, Université Lyon 1, Université de Saint-Etienne, Lyon, France
[5]Laboratory for Research on Learning and Development, LEAD – CNRS UMR5022, Université de Bourgogne, Dijon, France
[6]Cochlear Implant Center Northern Netherlands, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands

ABSTRACT

Although some cochlear implant (CI) listeners enjoy music, for others, music appreciation remains limited. Numerous studies have observed that perceiving musical emotions, i.e., the emotion depicted by a musical excerpt, was hindered due to limitations of electric hearing. Using acoustic vocoders to simulate some aspects of CI stimulation, we previously showed that arousal (e.g., joy vs. serenity), typically conveyed by rhythm information, was well transmitted whereas valence (e.g., sadness vs. serenity), typically conveyed by tonal-pitch information, was only poorly transmitted. The aim of the present study was to assess how CI group patterns would compare to what was predicted from vocoders, and whether individual results could be explained by performance in psychophysical tasks related to rhythm and pitch perception. Preliminary results indicate that the categorization pattern of CI listeners was similar to normal-hearing participants with vocoder simulation. Variability in arousal perception did not correlate with performance in a beat-alignment task that evaluated rhythm perception. However, valence perception correlated with voice-pitch discrimination thresholds. These results provide further insight in music emotion perception in CI listeners, and especially how valence transmission might be improved by a decreased discrimination threshold for pitch.

Keywords: *music emotion perception, arousal, valence, cochlear implants, individual differences.*

1. INTRODUCTION

Cochlear implant (CI) users report music enjoyment to be important to their quality of life [1], yet for numerous CI users music appreciation remains limited [2]. Emotion perception in music is one important contributing factor to its appreciation [3], and music emotion perception is reportedly challenging in CI users [4]. Music-evoked or -perceived emotion can be described along arousal (exciting, relaxing) and valence (positive, negative) dimensions [5], where, typically, rhythmic cues convey emotional arousal and tonal-pitch relationships convey emotional valence [6].

Regarding the perception of music emotion, CI users have been observed to more accurately perceive arousal than

[1]****Corresponding author**: *e.e.harding@rug.nl*

valence [7], [8], cf. [9]. This finding is intuitive as rhythmic cues are more available in the acoustic envelopes of sound, which are better conveyed by the implant [10] than is pitch information in the frequency spectrum of sound. This pattern was recently replicated in normal-hearing (NH) participants who performed a music emotion categorization task with vocoded excerpts that intended to approximate some aspects of limitations related to CI hearing and electric stimulation [11].
Different sensory and cognitive mechanisms contribute to rhythm and pitch perception. Therefore, the current study with CI users further assessed whether scores on a rhythm discrimination test correlated with how well arousal features were perceived, and whether fundamental frequency (pitch) discrimination thresholds correlated with how well emotional valence features were perceived.

2. MATERIALS AND METHODS

2.1 Participants

Participants were 25 adult CI users (mean age: 61.28 years, SD: 15.28, age range 21-82 years) who were implanted at least one year prior to testing. Participants considered themselves non-musicians and had less than three years of formal musical training, which ended more than two years before participation.

2.2 Materials

Experimental stimuli were 40 musical excerpts corresponding to four emotions (10 per emotion): joy, fear, serenity, and sadness. Stimuli were modifications from the original source [5], and fully described in [11] corresponding to the non-vocoded condition.
The rhythm test was the Beat-Based Alignment test [12], implemented as in [13]. Stimuli consisted of 32 rhythms presented as a pure tone in one of 6 frequencies (294, 353, 411, 470, 528, and 587 Hz, selected at random); half were "simple" rhythms (strong beat-based metrical structure and generally easier to discriminate) and half were "complex" rhythms (weaker metrical structure due to syncopation and generally more challenging to discriminate).
The voice pitch test was the Just-Noticeable Difference test (JND, [14]) for voice discrimination where voice stimuli were changed along pitch (fundamental frequency; F0) or speaker size (measured in vocal tract length; VTL) dimensions. Speech samples were presented with one odd sample that differs from the original voice in VTL and/or F0 values presented in two alternatives in an adaptive procedure. The JNDs were expressed as representing a change in semitones along F0 and VTL parameters.

2.3 Procedure

Participants were seated in a comfortable chair in front of a Microsoft Surface touchscreen tablet whose sound was outputted via two Logitech loudspeakers. The loudspeakers were placed to the left and right of the tablet.

Music emotion experiment: Participants received written instructions to listen to musical excerpts and choose the corresponding category of emotion that was presented in the excerpt. The task was embedded in an illustrated story about aliens who want to understand earth's music. The examiner answered any questions, and 8 practice trials (two trials of each emotion category) were given before the experiment to make sure they understood the task.
Rhythm test: Participants evaluated whether a target rhythm was the same or different from two preceding identical rhythms.
Pitch test: Participants identified which of three pseudoword utterances were spoken by a voice different from the other two. The 'different' voice could occur in any position (1,2, or 3).

3. ANALYSIS

The emotion categorization raw data was compiled in a confusion matrix where a 'perfect' score would be represented by a diagonal line, where presented categories (x-axis) were always the same as responded categories (y-axis). Confusion matrices were followed up with a Feature Information Transmission Analysis (FITA; [15]). In the FITA, considered features were emotional arousal and valence classes for each category: emotions were assigned positive or negative valence and low or high arousal. The outcome (T_{rel}) was a continuous variable between 0 and 1 and represented the relative quantity of transmitted information in proportion to the total available information. T_{rel} was logit-transformed and entered into an ANOVA with factor Feature (Arousal, Valence). Additionally, we further conducted Pearson's correlations to assess whether rhythm and fundamental frequency (F0) discrimination scores correlated with arousal – and valence feature transmission, respectively. Analyses were implemented in R [16].

4. PRELIMINARY RESULTS

4.1 Proportion correct

Participants were able to categorize emotions above chance-level or 0.25 proportion-correct (mean, SD): Joy = 0.60 (0.17); Fear = 0.54 (0.18); Serenity = 0.45 0.15); Sadness = 0.48 (0.14).

4.2 Confusion matrices

Raw data is presented in the form of confusion matrices (Figure 1). High arousal emotions (joy and fear) were confused with each other, as were low arousal emotions confused with each other (serenity and sadness). High- or low-valence emotions (joy and serenity; fear and sadness) were seldom confused.

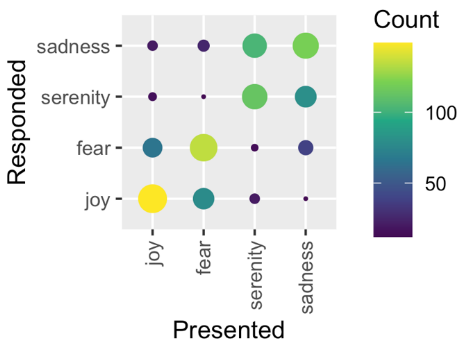

Figure 1. Confusion matrices for emotion categorization. Categorization was confused across valence conditions but within arousal conditions. The size and color of each dot is proportional to the number of relevant responses. The presented categories are listed on the x-axis, the responded categories are listed on the y-axis.

4.3 Feature Information Transmission Analysis (FITA)

The FITA analysis statistically addresses the visual pattern found in the above confusion matrix. Features arousal and valence were assessed [15]. A one-way ANOVA with factor Feature (Arousal, Valence) showed a significant difference in the feature transmission: Arousal was conveyed more efficiently than valence ($F(1,48) = 85.05$; $p < .001$)

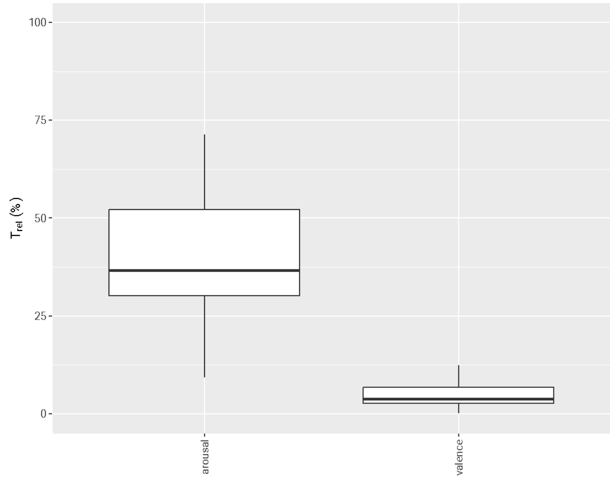

Figure 2. Arousal and valence feature transmission. When CI users categorized music emotion, the musical excerpts conveyed emotional arousal more than emotional valence. T_{rel} on the y-axis is the outcome of FITA (see Analysis section).

4.4 Correlations

Two correlations were performed to investigate the potential contribution of participants' abilities as measured in rhythm or pitch tasks to arousal and valence perception of the music emotion stimuli. While the rhythm-arousal correlation was not-significant ($R = .11$, $p = 0.60$, N=25), F0 perception correlated significantly with valence feature transmission ($R = .43$; $p = .032$).

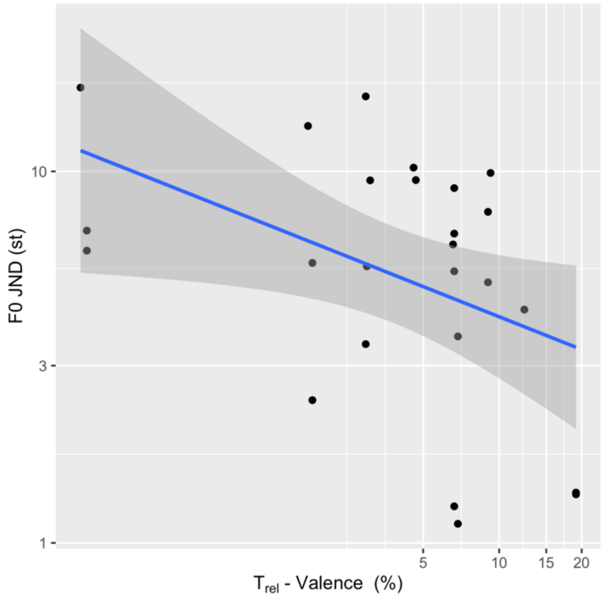

Figure 3. Correlation between emotional valence transmission and fundamental frequency (F0) perception in speech. In CI users, the perception of emotional valence in music was higher for implant users whose thresholds for voice pitch discrimination in speech was lower (lower values indicate better sensitivity to voice pitch cues).

5. CONCLUSIONS

These preliminary findings lead us to two conclusions, if confirmed with additional data collection. First, approximating CI hearing with a spread of vocoding parameters (improved or reduced quality of temporal and spectral information) tested with normal-hearing listeners [3] can capture actual CI performance for music emotion categorization with reasonable accuracy. Thus this method shows good potential for future music emotion paradigms as a way to both prepare tests for CI-user participants and to make estimates about the range of their performance. Second, the weak but significant correlation between F0 discrimination thresholds and emotional valence perception suggests that in CI users, perceptual mechanisms related to pitch perception thresholds in voice may be associated with the perception of tonal relationships required to perceived emotional valence in music, and this relationship warrants further investigation. Additional data collection for this study is planned.

6. ACKNOWLEDGMENTS

We would like to thank Dr. Christina Fuller for support in conceptual planning of this design and Prof. Emmanuel Bigand for kindly sharing the musical excerpt materials.

7. REFERENCES

[1] C. Fuller, R. Free, B. Maat, and D. Başkent, "Self-reported music perception is related to quality of life and self-reported hearing abilities in cochlear implant users," *Cochlear Implants International*, vol. 23, no. 1, pp. 1–10, Sep. 2021, doi: 10.1080/14670100.2021.1948716.

[2] C. Fuller, D. Başkent, and R. Free, "Early Deafened, Late Implanted Cochlear Implant Users Appreciate Music More Than and Identify Music as Well as Postlingual Users," *Frontiers in Neuroscience*, vol. 13, no. October, 2019, doi: 10.3389/fnins.2019.01050.

[3] O. Grewe, F. Nagel, R. Kopiez, and E. Altenmüller, "How Does Music Arouse 'Chills'?," *Annals of the New York Academy of Sciences*, vol. 2060, no. 1, pp. 446–449, 2005.

[4] S. Paquette, G. D. Ahmed, M. V. Goffi-Gomez, A. C. H. Hoshino, I. Peretz, and A. Lehmann, "Musical and vocal emotion perception for cochlear implants users," *Hearing Research*, vol. 370, pp. 272–282, 2018, doi: 10.1016/j.heares.2018.08.009.

[5] E. Bigand, S. Vieillard, F. Madurell, J. Marozeau, and A. Dacquet, "Multidimensional scaling of emotional responses to music: The effect of musical expertise and of the duration of the excerpts," *Cognition & Emotion*, vol. 19, no. 8, pp. 1113–1139, Dec. 2005, doi: 10.1080/02699930500204250.

[6] G. Ilie and W. F. Thompson, "A Comparison of Acoustic Cues in Music and Speech for Three Dimensions of Affect," *Music Perception*, vol. 23, no. 4, pp. 319–330, Apr. 2006, doi: 10.1525/mp.2006.23.4.319.

[7] K. L. D'Onofrio, M. Caldwell, C. Limb, S. Smith, D. M. Kessler, and R. H. Gifford, "Musical Emotion Perception in Bimodal Patients: Relative Weighting of Musical Mode and Tempo Cues," *Frontiers in Neuroscience*, vol. 14, p. 114, 2020, doi: 10.3389/fnins.2020.00114.

[8] S. Giannantonio, M. J. Polonenko, B. C. Papsin, G. Paludetti, and K. A. Gordon, "Experience changes how emotion in music is judged: Evidence from children listening with bilateral cochlear implants, bimodal

devices, and normal hearing," *PLoS ONE*, vol. 10, no. 8, pp. 1–29, 2015, doi: 10.1371/journal.pone.0136685.

[9] E. Ambert-Dahan, A. L. Giraud, O. Sterkers, and S. Samson, "Judgment of musical emotions after cochlear implantation in adults with progressive deafness," *Frontiers in Psychology*, vol. 6, no. MAR, pp. 1–11, 2015, doi: 10.3389/fpsyg.2015.00181.

[10] R. V. Shannon, "Temporal modulation transfer functions in patients with cochlear implants," *The Journal of the Acoustical Society of America*, vol. 91, no. 4, pp. 2156–2164, Apr. 1992, doi: 10.1121/1.403807.

[11] E. E. Harding *et al.*, "Musical Emotion Categorization with Vocoders of Varying Temporal and Spectral Content," *Trends in Hearing*, vol. 27, p. 1-19. Jan. 2023, doi: 10.1177/23312165221141142.

[12] M. Niarchou, J. F. Sathirapongsasuti, and N. Jacoby, "Unravelling the genetic architecture of musical rhythm," *BiorXiv*, pp. 1–46, 2019, doi: 10.1101/836197.

[13] A. Fiveash, S. D. Bella, E. Bigand, R. L. Gordon, and B. Tillmann, "You got rhythm, or more: The multidimensionality of rhythmic abilities," *Atten Percept Psychophys*, vol. 84, no. 4, pp. 1370–1392, May 2022, doi: 10.3758/s13414-022-02487-2.

[14] L. Nagels, E. Gaudrain, D. Vickers, P. Hendriks, and D. Başkent, "Development of voice perception is dissociated across gender cues in school-age children," *Sci Rep*, vol. 10, no. 1, p. 5074, Mar. 2020, doi: 10.1038/s41598-020-61732-6.

[15] G. A. Miller and P. E. Nicely, "An Analysis of Perceptual Confusions Among Some English Consonants," *The Journal of the Acoustical Society of America*, vol. 27, no. 2, pp. 338–352, Mar. 1955, doi: 10.1121/1.1907526.

[16] R Core Team, "R: A language and environment for statistical computing." R Foundation for Statistical Computing, Vienna, Austria, 2020. [Online]. Available: https://www.r-project.org/