JASA EXPRESS LETTERS ARTICLE asa.scitation.org/journal/jel

Hearing aid evaluation for music: Accounting for acoustical variability of music stimuli

Christophe Lesimple,[1] **Volker Kuehnel,**[1] **and Kai Siedenburg**[2,3,a)]
[1]*Sonova AG, Stäfa, Switzerland*
[2]*Signal Processing and Speech Communication Laboratory, Graz University of Technology, Graz, Austria*
[3]*Department of Medical Physics and Acoustics, Carl von Ossietzky Universität Oldenburg, Oldenburg, Germany*
christophe.lesimple@sonova.com, volker.kuehnel@sonova.com, kai.siedenburg@tugraz.at

Abstract: Music is an important signal class for hearing aids, and musical genre is often used as a descriptor for stimulus selection. However, little research has systematically investigated the acoustical properties of musical genres with respect to hearing aid amplification. Here, extracts from a combination of two comprehensive music databases were acoustically analyzed. Considerable overlap in acoustic descriptor space between genres emerged. By simulating hearing aid processing, it was shown that effects of amplification regarding dynamic range compression and spectral weighting differed across musical genres, underlining the critical role of systematic stimulus selection for research on music and hearing aids.

[Editor: Vasileios Chatziioannou] https://doi.org/10.1121/10.0028397

Received: 25 June 2024 **Accepted:** 9 August 2024 **Published Online:** 5 September 2024

1. Introduction

Participation in musical activities can promote social cohesion, can contribute to an individual's overall well-being (Bullack *et al.*, 2018; Clift and Hancox, 2010), and is thought to offer numerous social, psychological, and cognitive benefits. For individuals with hearing impairments, these benefits can be particularly significant, as music can help alleviate feelings of isolation and improve self-esteem (Darrow, 2006). However, the enjoyment and benefits of music can be compromised by the limitations of current hearing aid technology, which is primarily optimized for speech perception (Chasin and Russo, 2004). Therefore, improving music perception through hearing aids has potential to significantly enhance the quality of life of hearing aid users.

A special music program in most hearing aids typically involves modifications of the standard signal processing components to better accommodate the unique characteristics of music. First, amplification rules are often adjusted to enhance perception of lower frequencies while preserving audibility over the widest possible bandwidth (Moore, 2016; Moore *et al.*, 2016; Vaisberg *et al.*, 2021). Second, dynamic range compression (DRC), which reduces the range of output levels between the weakest and strongest sounds, is often minimized or turned off in the music program because excessive DRC can distort the musical signal and reduce sound quality (Arehart *et al.*, 2011; Croghan *et al.*, 2014; Kirchberger and Russo, 2016a; Madsen *et al.*, 2015; Vaisberg *et al.*, 2021). However, results from experiments using recorded stimuli might not apply to listening situations with a larger dynamic range where audibility might be challenged with more linear amplification. Third, frequency lowering, a technique used to make high-frequency sounds audible by shifting them to lower frequencies, may also be specifically adapted to improve the quality of music perception (Kirchberger and Russo, 2016b). Last, adaptive features such as feedback cancellation or noise reduction are often deactivated or reduced in the music program. The latter feature, while beneficial for speech understanding in noisy environments, can inadvertently alter the musical signal and degrade the listening experience (Greasley *et al.*, 2020; Kim *et al.*, 2020; Madsen and Moore, 2014).

Research on hearing aid signal processing and its impact on music perception often employs different music genres as part of the evaluation process. This approach is based on the assumption that different music genres, with their unique acoustic characteristics, can interact in non-systematic ways with various signal processing changes, thereby influencing the perceived audio quality (Arehart *et al.*, 2011; Kirchberger and Russo, 2016a; Moore and Sęk, 2016). The interaction between music genre and hearing aid processing is complex and can vary, depending on the specific characteristics of the music and the type of signal processing applied (Madsen and Moore, 2014; Vaisberg *et al.*, 2021). For instance, DRC (Croghan *et al.*, 2014) or frequency lowering (Kirchberger and Russo, 2016b) might be perceived differently, depending on the music genre. This non-systematic interaction underscores the complexity of optimizing hearing aid technology for music listening.

[a)]Author to whom correspondence should be addressed.

The acoustical properties of music samples used for audio quality evaluation, particularly dynamic range and long-term spectrum, are key considerations in the cited studies. Kirchberger and Russo (2016a) reported that samples from more classical genre, such as opera and orchestra, typically have a wider dynamic range than samples from more modern music, like pop, rock, or rap. The long-term spectrum is also flatter for music with percussive sounds present in more modern music productions (Arehart *et al.*, 2011; Davies-Venn *et al.*, 2007). The dynamic range and the spectrum of the input signal play a critical role when evaluating different DRC approaches and might also explain some interactions between the different tested signal processing conditions and the tested samples (Croghan *et al.*, 2014; Davies-Venn *et al.*, 2007; Moore and Sęk, 2016; Vaisberg *et al.*, 2021). So far, this interaction has not been systematically evaluated. Moreover, the process by which musical genre and test stimuli are selected is not always transparent in the literature.

Music genres (e.g., pop, rock, classical orchestral, or choir) reflect certain shared conventions or characteristics among pieces of music, such as commonalities in form, style, or subject matter (Fabbri, 1981). These characteristics can encompass a wide range of elements, including melody, rhythm, harmony, lyrics, instrumentation, and production techniques (Tagg, 2012). However, defining music genres is not always straightforward. The boundaries between genres and subgenres can often be fluid and subjective and may be influenced by cultural or historical trends (Brackett, 2023). While some genres are easily distinguishable by their unique characteristics, others may overlap, making it challenging to categorize a piece of music definitively into a single genre (Negus, 2013). This complexity reflects the rich diversity and evolving nature of music (Mauch *et al.*, 2015).

While music genre serves as a useful categorization tool in hearing aid research, it presents a significant challenge: A single music sample used for testing, although defined by its genre, may not be representative of the entire genre. This is due to the inherent variability within each genre, a factor that is neither well understood nor extensively studied. Consequently, generalizing the findings of audio quality evaluations based on a single sample to an entire genre becomes problematic. This leads us to two research questions: (1) How can we generalize the findings of audio quality evaluations to an entire genre given the within-genre variability? (2) Is the effect of hearing aid amplification homogeneous both between genres and within a single genre? Addressing these questions would improve our understanding of music perception with hearing aids and help to optimize hearing aid technology for diverse music listening experiences.

In the present study, we assessed the acoustical variability of music stimuli using an approach based on elementary spectral and level-based audio descriptors. We considered variability between and within genres as well as between different time windows within pieces. Further, we studied how hearing aid processing using fast-acting and slow-acting DRC is affected by this variability, in terms of descriptors assessing dynamic range, spectral centroid (SC), and ratio of harmonic to percussive energy.

2. Methods

2.1 Audio database, sample selection, loudness normalization

A database of recorded music was compiled based on previous research projects (Gerdes and Siedenburg, 2023; Kirchberger and Russo, 2016a), together yielding a repository of 1403 pieces of music. Music genre labels were used as stated in the reference papers and were arranged into 12 classes, representing a range of different types of Western music, namely chamber, orchestra, opera, choir, piano, jazz, country, schlager (a Central European pop music genre usually with German lyrics), rap, pop, rock, and metal.

Each entire piece was first normalized in loudness relative to full scale according to the ITU-R BS.1770 standard (International Telecommunication Union, 2003) and then adjusted to be played back at 70 dB sound pressure level (SPL) for the hearing aid input level. This approach approximately matches the average preferred listening level for hearing aid users (Croghan *et al.*, 2016) and allows the level to vary between extracted samples, as would naturally be the case when listening to a piece with loud and soft passages. Three 20 s samples were randomly extracted from each third of the piece to add some within-piece variability to the data set, reflecting the possibility that a single piece may feature soft and loud excerpts or different instrumentations. This procedure produced a total of 4209 music samples of 20 s duration. The acoustical analysis was performed on the normalized raw samples processed by a hearing aid simulator with 0 dB linear insertion gain and with compressive amplification.

2.2 Computation of audio descriptors

Each sample was available in stereo wave file format. Audio descriptors were extracted for both left and right channels and then averaged. The descriptors focused primarily on energy and spectral information, as these are the main aspects affected by the gain model under investigation (Kirchberger and Russo, 2016a).

The root-mean square (RMS) energy of the signal was calculated with a sliding rectangular window with 93 ms duration and a step size of 46 ms (Müller, 2015). The peak-to-average ratio was also selected to provide additional information about the dynamic range of the signal (Chasin and Russo, 2004). It was expected that DRC would increase the median root-mean-square (RMS) value and reduce the standard deviation (SD) of the RMS values (RMS SD).

Spectral information was extracted using statistical moments of the spectral distribution (Caetano *et al.*, 2019). The SC was calculated as the center of gravity of the frequencies present in the sound (with amplitude magnitudes as weights) (Tzanetakis and Cook, 2002), corresponding to the perceived "brightness" of a musical mix (Siedenburg *et al.*, 2021). The spectral bandwidth, an amplitude-weighted SD of the spectrum, provides a measure of the dispersion of the energy in the spectrum around its mean frequency (Müller, 2015). SC and spectral bandwidth were computed from time-frequency representations based on the center frequencies of a GAMMATONE FILTERBANK (MathWorks, Natick, MA) evenly spaced on the equivalent rectangular bandwidth (ERB) scale (Patterson *et al.*, 1995). The median, SD, and skewness of the audio descriptor distributions were extracted to summarize spectral information for each sample.

Three additional audio descriptors were extracted as they could provide complementary indirect information about energy and spectral content: zero-crossing rate, harmonic-to-percussive ratio (H2P), and spectral flatness via Wiener entropy (Müller, 2015). Regarding the latter descriptor, a sound is said to be more "noisy" or "atonal" when it has higher spectral flatness and more "tonal" when it has lower spectral flatness. Broadband zero-crossing rate is the rate of a signal's sign change and has been successfully used for music genre classification (Yi *et al.*, 2021). Harmonic-percussive source separation based on median filtering was used to estimate the energy ratio between extracted harmonic and percussive signals (H2P). Filtering along the time axis will retain the harmonic or pitched signals, and filtering along the frequency axis will extract the percussive components (Müller, 2015). The central tendency and SD of these audio descriptors were computed. These descriptors were extracted using the LIBROSA library (McFee *et al.*, 2015).

2.3 Hearing aid amplification

Hearing aid amplification was applied with an audio plugin (AUDIO PLUGIN GENERATOR; dlab, Winterthur, Switzerland) using batch processing from the digital audio workstation Adobe Audition (Adobe Systems, San Jose, CA). Initial processing using a 0 dB flat insertion gain was used to generate the real-ear unaided response as a reference. The real-ear aided response was generated with the fast- and slow-acting DRC settings to apply the amplification defined by the NAL-NL2 fitting rationale (Keidser *et al.*, 2011) for a moderate-to-severe N4 hearing loss (Bisgaard *et al.*, 2010). The effect of amplification was then computed as the difference between the descriptors for the aided and the unaided real-ear responses to ensure identical bandwidth and to account for the real-ear unaided gain.

DRC was applied in 20 bands approximating the Bark scale from 100 Hz to 10 kHz. Compressive amplification had an average compression ratio of 2.5. It was applied separately as fast-acting DRC with an attack time (AT) of 10 ms and release time (RT) of 60 ms and as slow-acting DRC with an AT of 1 s and RT of 8 s (Chen *et al.*, 2021). With fast-acting DRC, the gain is adjusted with a finer temporal resolution than with slow-acting DRC in order to make weaker parts of the input signal more audible and to prevent loudness discomfort from sudden loud sounds. However, this might introduce unwanted artefacts, compromising audio quality as fast-acting DRC is known to flatten the temporal envelope. It is also expected that the spectral envelope is flattened if level estimation is performed independently in all compression bands, reducing the contrast between high and low energy parts of the spectrum (Croghan *et al.*, 2014; Plomp, 1988).

3. Results

3.1 Between-genre variability

Variability and correlations between the 16 audio descriptors, extracted for all the unprocessed samples, were summarized in a principal component analysis (PCA). The median and SD were obtained from the distributions of the RMS, SC, spectral bandwidth, spectral flatness, zero-crossing rate, and H2P ratio. Skewness of the RMS, SC, and spectral bandwidth and peak-to-average value were also calculated. All the data were centered and scaled to unit variance before computing the PCA. The first four dimensions from the PCA were retained as their eigenvalues were higher than 1. These dimensions explained 80.4% of the cumulative variance in the data. The first dimension explained 44.5% and the second dimension 18.9% of the total variance. Individual observations grouped by genre as well as the loadings of the original variables on the first two principal components are shown in Fig. 1, including genre-specific 95% confidence ellipses.

Considering the contribution of individual audio descriptors, the first dimension received strong contributions of the median SC (11.4%), spectral bandwidth (11.1%), spectral flatness (11.0%), and H2P ratio (9.8%). There were negative correlations between H2P ratio and median SC ($r = -0.72$), bandwidth ($r = -0.70$), and flatness ($r = -0.73$), suggesting that percussive components have a flatter and wider spectrum. This matches the mean genre distribution in the two-dimensional (2D) projection, ranging from piano, orchestra, choir, chamber, and opera, with the lowest values on dimension 1 (corresponding to a high H2P ratio and low SCs, bandwidth, and flatness), to jazz and country, in the middle range of dimension 1, up to rap, metal, schlager, pop, and rock, with the highest values on dimension 1.

The second dimension received strong contributions of RMS SD (13.5%), SC SD (12.4%), and the peak-to-average ratio (11.4%). The RMS SD and peak-to-average ratio were highly correlated ($r = 0.86$), both reflecting the dynamic range of the signal. Interpreting the exact acoustic meaning of this dimension is not straightforward. However, within the more "modern" genres, the results suggest that there is more variability in the audio descriptors in decreasing order for the rap, pop rock, schlager, and metal genres. Overall, the confidence ellipses suggest that there is large within-genre variability, leading to a large overlap of audio descriptor distributions between genres. Only the ellipse for the piano genre was clearly separated from the ellipses for the rap and schlager genres. This indicates that the gross acoustical properties of most popular genres overlap to a substantial extent.

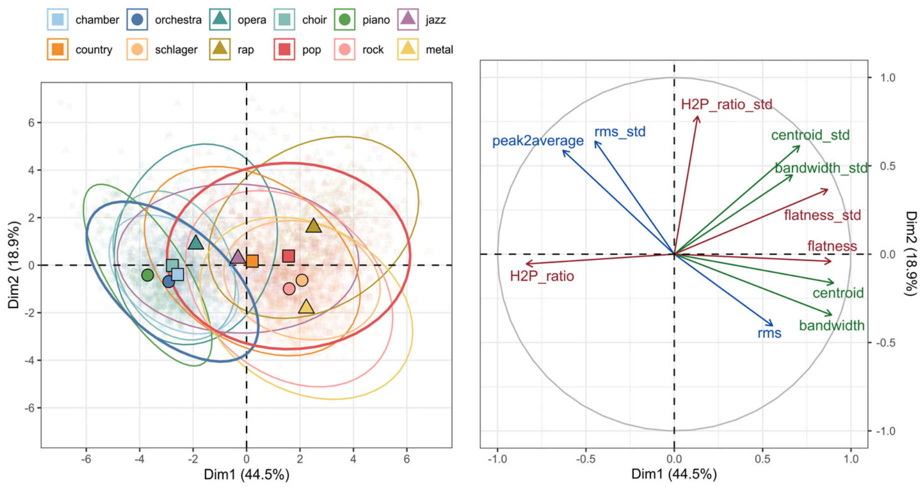

Fig. 1. PCA with music genre as grouping variable (left panel) and selected relevant variables (contribution higher than 10%) (right panel) projected on the two first dimensions (Dim1 and Dim2). Confidence ellipses show the 95% confidence interval for the bivariate distributions. The most often used genres in the referenced papers, i.e., classical orchestral and pop music, are highlighted to illustrate the overlap. rms, root mean square; std, standard deviation.

3.2 Within-piece variability

Within-piece variability for four selected recordings, the excerpts of which were closest to the genre mean in the PCA space, is illustrated in Fig. 2 for two orchestral pieces and two popular genre pieces. Projected audio descriptor distribution on the PCA dimensions for the first classical piece (Mozart, *Symphonie No. 1*) was located in one specific domain of the descriptor space, namely towards the central portion of the space. However, the second classical piece (Smetana, *Hakon Jarl*) gave a distribution of values that filled more than half of the confidence ellipse of the orchestra genre in addition to a specific trajectory beyond the ellipse boundaries. Note that this piece is also extraordinarily diverse from a musical point of view, so it is no surprise that the audio descriptor distribution falls outside the regions of relatively high density for the orchestra genre. A similar, albeit a little less drastic situation occurred for the two popular genre pieces. Whereas the first piece (Sheryl Crow, *All I Wanna Do*) exhibited a fairly focused distribution of values within the cluster for popular music, audio descriptor values of the second piece (Lady Gaga, *Just Dance*) were more widely spread in the space and fell partly outside the confidence ellipse.

Based on this global analysis of a large set of pieces from different musical genres as well as "microscopic" analysis of the excerpts from single pieces, we conclude that isolated excerpts of randomly drawn pieces are often not representative of the acoustical properties of a specific musical genre nor representative of the piece itself. A pragmatic, statistically based recommendation from these findings is to include samples and genre variables as random effects in the statistical evaluation (Barr *et al.*, 2013), which may help to generalize findings beyond the selected test material.

3.3 Effect of amplification by genre

Figure 3 shows a 2D density plot of the distribution of the RMS SD and SC descriptors on the left and H2P ratio and SC SD on the right for the unaided condition in the top row. The median effect of hearing aid amplification (i.e., the

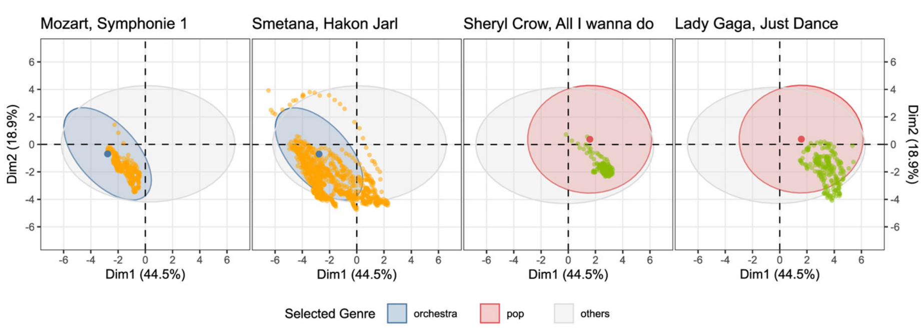

Fig. 2. Samples from single classical (two leftmost panels) and pop (two rightmost panels) pieces obtained from 20 s samples with a 1 s sliding window. The reference ellipses for pop and classical genres are taken from Fig. 1. Examples with more homogeneous samples within a single piece and with more heterogeneity between samples are shown for both genres.

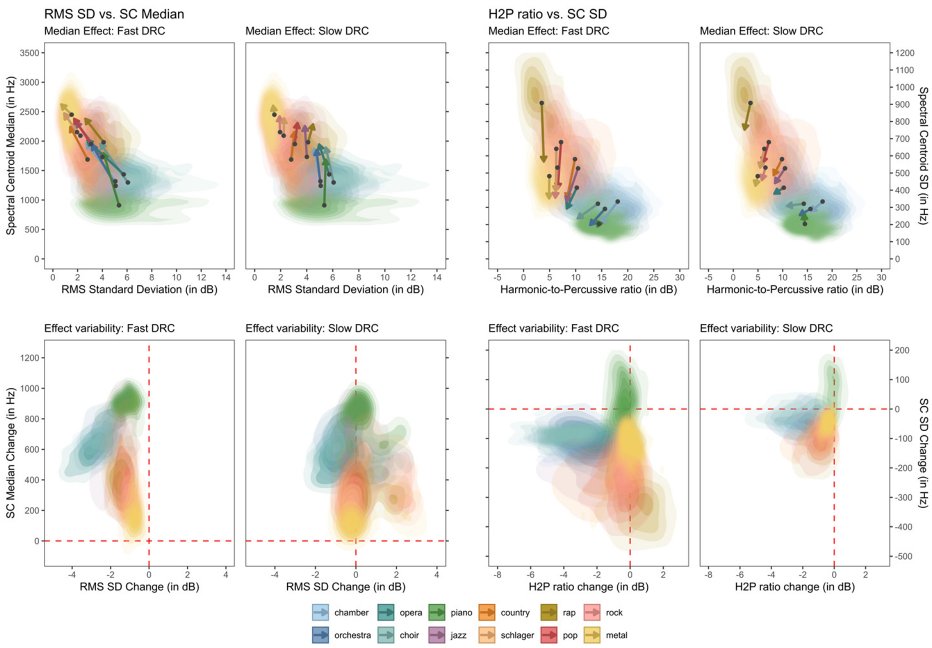

Fig. 3. The median amplification effect (aided/unaided) for each genre for fast and slow DRC is shown in the top row via arrows. The unaided reference distribution is represented by a two-dimensional density plot. Variability of the amplification effect for each genre is plotted as a bivariate distribution in the bottom row. The reference lines in red represent no effect. The relationship between the RMS SD and SC median is shown on the left, and the relationship between H2P ratio and SC SD is shown on the right.

difference of descriptor values in the condition with 0 dB insertion gain minus descriptor values in the aided condition) is represented via arrows (top row), and its variability is shown with a bivariate distribution of the amplification effect on the bottom row.

For the relationship between the RMS SD and SC median (left part), there are no clear separating boundaries between any of the genres, with smooth transitions in audio descriptor values. Genres using solely acoustical instruments (piano, orchestral) had lower SC values and generally higher RMS SD values. This may partly reflect less drastic use of compression in these recordings. Notably, the effects of amplification on SC values are much more pronounced for these more acoustic genres. Whereas there is a median change of SC by about 1 kHz for piano excerpts, there is little amplification effect for genres with higher unaided values. As expected, the effects of slow vs fast DRC differ mainly regarding the change in RMS SD. Whereas fast DRC reduces the RMS SD of chamber and orchestral music excerpts by up to 3 dB, the reduction is only about 1 dB for genres such as pop, metal, or piano music. For slow DRC, there was no substantial change in RMS SD, although excerpts of classical pieces were somewhat compressed.

For the relationship between the H2P ratio and SC SD (right part), the unaided bivariate distribution suggests that genres with acoustic instruments (e.g., piano, orchestra) have a higher H2P ratio and lower SC SD, indicating that these genres generally comprise sounds with weaker percussive components. Amplification tends to provide a stronger reduction of SC SD for genres with initially high values such as rap or pop. Furthermore, the magnitude of the reduction via amplification is more pronounced for fast DRC than for slow DRC. The piano genre has a slightly different behavior for the SC SD, with no effect for the fast DRC and a slight median increase in 60 Hz for the slow DRC. In contrast, no clear changes in H2P ratio were observed for genres such as metal (–0.3 dB) or rap (–0.4 dB). Overall, both the RMS SD vs SC plots and the H2P vs SC SD plots indicated effects of DRC that were different across the array of genres.

4. Summary and discussion

To evaluate hearing aids and any other hearing technology for music, ecologically valid test stimuli should be used. Our results indicate, however, that the acoustical properties of music excerpts defined solely by categories of musical genre may not be representative. The fact that a musical excerpt is labelled "jazz" does not mean that its acoustical properties will be different from excerpts from other genres. Therefore, there is the risk of over-interpreting results based on a small number of samples. Generally, our results imply that the description of the acoustical properties of test stimuli used for evaluating hearing aids might provide more relevant information than their genre label. A rationale for excerpt selection should be developed with regard to existing within-piece variability, for which the current analysis may serve as a starting point. Furthermore, the analysis of aided dynamic range and spectral energy distributions, reflecting the effect of the two critical hearing aid functions of DRC and frequency response, suggests that hearing aids differentially affect excerpts with different dynamic range and spectral properties. In sum, a more detailed acoustical description of test conditions is needed to enable the design of rigorous and reproducible listening experiments into hearing aids for music.

Acknowledgments

The authors thank Simon Jacobsen and Brian C. J. Moore for providing valuable comments on the manuscript. K.S. was funded by the Deutsche Forschungsgemeinschaft [DFG (German Research Foundation)]—Project ID No. 352015383957—SFB 1330 A6. K.S. was also supported by a Freigeist Fellowship from the Volkswagen Foundation.

Author Declarations

Conflict of Interest

We disclose that Christophe Lesimple and Volker Kuehnel are employees of Sonova AG.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

Arehart, K. H., Kates, J. M., and Anderson, M. C. (**2011**). "Effects of noise, nonlinear processing, and linear filtering on perceived music quality," Int. J. Audiol. **50**(3), 177–190.

Barr, D. J., Levy, R., Scheepers, C., and Tily, H. J. (**2013**). "Random effects structure for confirmatory hypothesis testing: Keep it maximal," J. Mem. Lang. **68**(3), 255–278.

Bisgaard, N., Vlaming, M. S., and Dahlquist, M. (**2010**). "Standard audiograms for the IEC 60118-15 measurement procedure," Trends Amplif. **14**(2), 113–120.

Brackett, D. (**2023**). *Interpreting Popular Music* (University of California Press, Oakland, CA).

Bullack, A., Gass, C., Nater, U. M., and Kreutz, G. (**2018**). "Psychobiological effects of choral singing on affective state, social connectedness, and stress: Influences of singing activity and time course," Front. Behav. Neurosci. **12**(223), 1–10.

Caetano, M., Saitis, C., and Siedenburg, K. (**2019**). "Audio content descriptors of timbre," in *Timbre: Acoustics, Perception, and Cognition*, edited by K. Siedenburg, C. Saitis, S. McAdams, A. N. Popper, and R. R. Fay (Springer, New York), pp. 297–333.

Chasin, M., and Russo, F. A. (**2004**). "Hearing aids and music," Trends Amplif. **8**(2), 35–47.

Chen, Y., Wong, L. L., Kuehnel, V., Qian, J., Voss, S. C., and Shangqiguo, W. (**2021**). "Can dual compression offer better Mandarin speech intelligibility and sound quality than fast-acting compression?," Trends Hear. **25**, 233121652199761.

Clift, S., and Hancox, G. (**2010**). "The significance of choral singing for sustaining psychological wellbeing: Findings from a survey of choristers in England, Australia and Germany," Music Performance Res. **3**, 79–96.

Croghan, N. B., Arehart, K. H., and Kates, J. M. (**2014**). "Music preferences with hearing aids: Effects of signal properties, compression settings, and listener characteristics," Ear Hear. **35**(5), e170–e184.

Croghan, N. B., Swanberg, A. M., Anderson, M. C., and Arehart, K. H. (**2016**). "Chosen listening levels for music with and without the use of hearing aids," Am. J. Audiol. **25**(3), 161–166.

Darrow, A.-A. (**2006**). "The role of music in deaf culture: Deaf students' perception of emotion in music," J. Music Ther. **43**(1), 2–15.

Davies-Venn, E., Souza, P., and Fabry, D. (**2007**). "Speech and music quality ratings for linear and nonlinear hearing aid circuitry," J. Am. Acad. Audiol. **18**(08), 688–699.

Fabbri, F. (**1981**). "A theory of musical genres: Two applications," in *Popular Music: Critical Concepts in Media and Cultural Studies*, edited by S. Frith (Routledge, London), Vol. 3, pp. 7–35.

Gerdes, K., and Siedenburg, K. (**2023**). "Lead-vocal level in recordings of popular music 1946–2020," JASA Express Lett. **3**(4), 043201.

Greasley, A., Crook, H., and Fulford, R. (**2020**). "Music listening and hearing aids: Perspectives from audiologists and their patients," Int. J. Audiol. **59**(9), 694–706.

International Telecommunication Union (**2003**). "Algorithms to measure audio programme loudness and true-peak audio level," Technical Report 3 (International Telecommunication Union, Geneva, Switzerland).

Keidser, G., Dillon, H., Flax, M., Ching, T., and Brewer, S. (**2011**). "The NAL-NL2 prescription procedure," Audiol. Res. **1**(e24), 88–90.

Kim, H. J., Lee, J. H., and Shim, H. J. (**2020**). "Effect of digital noise reduction of hearing aids on music and speech perception," J. Audiol. Otol. **24**(4), 180–190.

Kirchberger, M., and Russo, F. A. (**2016a**). "Dynamic range across music genres and the perception of dynamic compression in hearing-impaired listeners," Trends Hear. **20**, 233121651663054.

Kirchberger, M., and Russo, F. A. (**2016b**). "Harmonic frequency lowering effects on the perception of music detail and sound quality," Trends Hear. **20**, 233121651562613.

Madsen, S. M., and Moore, B. C. J. (**2014**). "Music and hearing aids," Trends Hear. **18**, 233121651455827–233121651455829.

Madsen, S. M., Stone, M. A., McKinney, M. F., Fitz, K., and Moore, B. C. J. (**2015**). "Effects of wide dynamic-range compression on the perceived clarity of individual musical instruments," J. Acoust. Soc. Am. **137**(4), 1867–1876.

Mauch, M., MacCallum, R. M., Levy, M., and Leroi, A. M. (**2015**). "The evolution of popular music: USA 1960–2010," R. Soc. Open Sci. **2**(5), 150081.

McFee, B., Raffel, C., Liang, D., Ellis, D. P., McVicar, M., Battenberg, E., and Nieto, O. (**2015**). "Librosa: Audio and music signal analysis in Python," in *Proceedings of the 14th Python in Science Conference (SciPy2015)*, Austin, TX, Vol. 8, pp. 18–24.

Moore, B. C. J. (**2016**). "Effects of sound-induced hearing loss and hearing aids on the perception of music," J. Audio Eng. Soc. **64**(3), 112–123.

Moore, B. C. J., Baer, T., Ives, D. T., Marriage, J., and Salorio-Corbetto, M. (**2016**). "Effects of modified hearing aid fittings on loudness and tone quality for different acoustic scenes," Ear Hear. **37**(4), 483–491.
Moore, B. C. J., and Sęk, A. (**2016**). "Preferred compression speed for speech and music and its relationship to sensitivity to temporal fine structure," Trends Hear. **20**, 233121651664048.
Müller, M. (**2015**). *Fundamentals of Music Processing: Audio, Analysis, Algorithms, Applications* (Springer, Heidelberg).
Negus, K. (**2013**). *Music Genres and Corporate Cultures* (Routledge, London).
Patterson, R. D., Allerhand, M. H., and Giguere, C. (**1995**). "Time-domain modeling of peripheral auditory processing: A modular architecture and a software platform," J. Acoust. Soc. Am. **98**(4), 1890–1894.
Plomp, R. (**1988**). "The negative effect of amplitude compression in multichannel hearing aids in the light of the modulation-transfer function," J. Acoust. Soc. Am. **83**(6), 2322–2327.
Siedenburg, K., Barg, F. M., and Schepker, H. (**2021**). "Adaptive auditory brightness perception," Sci. Rep. **11**(1), 21456.
Tagg, P. (**2012**). *Music's Meanings: A Modern Musicology for Non-Musos* (Mass Media's Scholar's Press, Larchmont, NY).
Tzanetakis, G., and Cook, P. (**2002**). "Musical genre classification of audio signals," IEEE Trans. Speech Audio Process. **10**(5), 293–302.
Vaisberg, J. M., Beaulac, S., Glista, D., Macpherson, E. A., and Scollie, S. D. (**2021**). "Perceived sound quality dimensions influencing frequency-gain shaping preferences for hearing aid-amplified speech and music," Trends Hear. **25**, 233121652198990.
Yi, Y., Zhu, X., Yue, Y., and Wang, W. (**2021**). "Music genre classification with LSTM based on time and frequency domain features," in *2021 IEEE 6th International Conference on Computer and Communication Systems (ICCCS)*, Chengdu, China (IEEE, New York), pp. 678–682.