forum acusticum 2026
8-12 September · Graz, Austria

EFFECTS OF HRTF AUGMENTATION ON PREDICTED SPATIAL RELEASE FROM MASKING IN MUSIC

Jack Webb[1], Christophe Lesimple[2], Volker Kuehnel[2], Lorenzo Picinali[1]

[1] *Dyson School of Design Engineering, Imperial College London, United Kingdom*
[2] *Sonova AG, Stäfa, Switzerland*

This paper is a revised version of the authors' submitted manuscript that was peer-reviewed by at least two anonymous reviewers.

Abstract

Separating individual musical instruments within a complex mixture of sounds poses a persistent challenge for listeners with hearing loss. Although spatial separation of sources improves speech recognition in this population, the potential benefits of spatial cue enhancement for music perception remain largely unexplored. This paper introduces a method to increase spatial cue salience through the augmentation of individual head-related transfer functions (HRTFs). Auditory model analyses indicate that augmented HRTFs may enhance the separability of musical instruments relative to individual HRTFs. Predicted benefits persist when moderate sensorineural hearing loss is modelled, though they are substantially reduced. Simulated hearing aid processing does not restore these benefits to normal-hearing levels.

Keywords: *music scene analysis, hrtf, augmentation, spatial processing, hearing loss*

1. Introduction

Music listening for individuals with hearing loss is complicated by both the signal distortions introduced by hearing aids and a generally reduced acuity in music perception tasks [1]. In particular, the perceptual separation of individual musical instruments in a mixture of sounds remains a significant hurdle [2]. In essence this is a problem of auditory scene analysis, the process by which the auditory system decomposes complex sound environments into perceptually meaningful streams [3]. The ability to perform auditory scene analysis is directly influenced by masking, whereby the threshold for hearing one sound is raised by the presence of a competing one. Masking can be reduced through several mechanisms. Energetic release arises from spectral and modulation differences between target and masker sounds [4]. Higher-level processes such as attention, memory, and structural processing can also provide informational release from masking [5]. In complex acoustic scenarios, the spatial separation of target and masker provides cues that support both types of masking release [6].

1.1. Spatial Release from Masking in Music

Spatial release from masking (SRM) provides sizeable advantages to speech intelligibility in situations with multiple overlapping talkers. Although more common in speech literature, we use SRM here to denote the general use of spatial cues to separate multiple auditory sources. An analogous task occurs in music scene analysis (MSA), defined as the ability to discern a single instrument from competing musical sounds. Studies have shown that MSA performance generally improves with fewer competing instruments and elevated target-to-masker level ratios [7], as well as when fast dynamic range compression is used in hearing aids [8], and when mixes have greater spectral domain sparsity [9]. However, the role of SRM in MSA remains elusive. Only one work has analysed the effect of spatial cue manipulations on MSA performance: [7] identified a small but inconsistent benefit of increased stereo width on MSA, induced by frequency-independent level differences between left and right headphone channels. As such, the role of more complex spatial auditory cues in facilitating the segregation of music sources remains unclear.

A natural extension is to consider more complete spatial representations based on head-related transfer functions (HRTFs). Spatial auditory perception relies on interaural differences in time (ITD) and level (ILD), as well as monaural spectral cues. HRTFs encapsulate these cues, characterising how sound is directionally filtered by an individual listener's morphology before it reaches the eardrum. Unlike intensity stereo panning, which relies solely on frequency-independent ILDs, binaural rendering using HRTFs provides the auditory system with all three types of spatial cue, motivating exploration of their role in MSA.

Corresponding author: *j.webb24@imperial.ac.uk..*

1.2. Augmentation of Spatial Cues

Speech intelligibility research has investigated the possibility of increasing SRM by enhancing spatial cues. One approach involves designing filters that transform binaural signals to achieve linear magnification of interaural time and level differences [10]. Such algorithms have been shown to improve speech intelligibility in normal-hearing listeners [11]. Low frequency and broadband ILD magnification have also improved speech intelligibility in bilateral cochlear implant users [12]. Other works have achieved similar magnification outcomes by augmenting the interaural cues contained within HRTFs; in [13], augmentation of individually measured HRTFs resulted in improved speech intelligibility. Augmented HRTF spectra have also produced improvements in vertical localisation accuracy [14]. Despite potential benefits, all of these approaches have yet to be evaluated in MSA contexts.

1.3. Effects of Hearing Loss

Sensorineural hearing loss is known to contribute to compromised spatial auditory processing, leading to generally reduced SRM benefit [15]. Such deficits likely disrupt auditory scene analysis in listeners with hearing loss, whether this manifests as increased thresholds to understand speech in noise [16], or increased thresholds to hear specific instruments within a mixture [17]. Understanding the potential effects of hearing loss is essential to evaluating the utility of spatial cue augmentation for music listening in assistive devices.

1.4. Aims

Given the limited understanding of SRM in MSA, the uncertain impact of augmented spatial cues on SRM in music listening, and the potential deficits that may be introduced by hearing loss, the aims of this study are threefold:

1. To outline a framework for systematically augmenting interaural cues within individual HRTFs.
2. To estimate, using computational auditory models, the relative contribution of natural and augmented spatial cues to SRM in music listening.
3. To assess the likely influence of sensorineural hearing loss on access to SRM in music.

As this study relies on computational models of SRM and hearing loss, results are interpreted as predictions that provide a basis for future behavioural validation.

2. Methods

2.1. HRTF Augmentation

In a departure from previous augmentation approaches, we propose to enhance spatial cues by scaling the direction-dependent spectral components of a HRTF. While principal component analysis (PCA) has been long used in HRTF reconstruction [18], its application to spatial cue augmentation remains mostly unexplored. We performed PCA on HRTFs drawn from 405 individuals within the SONICOM HRTF dataset [19].

Let h represent the time-domain head-related impulse response (HRIR) from an arbitrary source position, with corresponding HRTF H in the frequency domain. Firstly, the directional transfer function (DTF) is obtained from each HRTF by subtracting the common transfer function (CTF) - the mean log-magnitude response averaged across all spatial locations - from each HRTF. The DTF is then decomposed into mid and side components. For left and right ear DTF spectra $L[k]$ and $R[k]$,

$$M[k] = \tfrac{1}{2}\big(L[k] + R[k]\big), \qquad S[k] = \tfrac{1}{2}\big(L[k] - R[k]\big),$$

where S captures interaural spectral differences including frequency-dependent ILDs. PCA is performed on this side channel S over all horizontal plane source positions and listeners. Observations were stacked into a matrix and decomposed by singular value decomposition as $S = U\Sigma V^\top$, where V contains the spectral basis functions and $s_i = U_i\Sigma$ gives the PCA scores for observation i. To magnify direction-dependent interaural structure, we apply residual contrast in score space: for each listener, we compute the mean side-channel score vector $\boldsymbol{\mu}$, averaged across horizontal directions, and scale only the deviation from this mean

$$\hat{\mathbf{s}}_i = \alpha\big(\mathbf{s}_i - \boldsymbol{\mu}\big),$$

using the first six principal components. All other scores are set to zero. The augmented side spectrum is reconstructed as $\hat{S}_i[k] = \hat{\mathbf{s}}_i V^\top$, and the augmented DTFs are rebuilt from the unchanged M component:

$$\hat{L}_i[k] = M_i[k] + \hat{S}_i[k], \qquad \hat{R}_i[k] = M_i[k] - \hat{S}_i[k].$$

The listener-specific CTF is then added back to obtain the augmented log-magnitude HRTF. The final augmented HRTF $\widehat{H}_i$ is obtained by exponentiating the augmented magnitude and restoring the original phase ($\angle H_i$). To minimise colouration, only the contralateral ear is augmented in each HRTF, and midline source positions are left unaltered. This formulation magnifies frequency-dependent ILDs while preserving broader monaural structure (Fig. 1).

2.2. Spatial Release from Masking Analysis

2.2.1. Stimuli

Numerical analysis was conducted to estimate the SRM benefit of different spatial cue combinations. A large dataset of musical samples was constructed using multitracks drawn from the Musiclarity and MedleyDB datasets [20], [21]. Each sample was 2 seconds in duration, and consisted of a single target instrument plus three masking instruments. The samples were each processed through seven rendering conditions: *Diotic*, fully diotic mix; *ILD*, frequency-

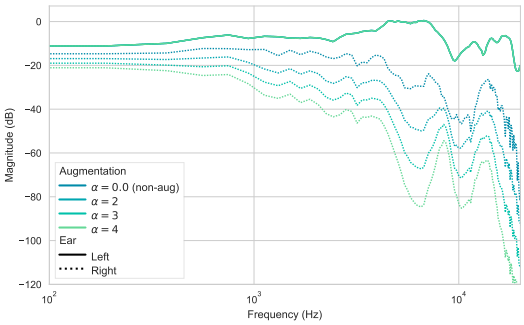

Figure 1: Example individual and augmented HRTFs for a source at 90° azimuth on the horizontal plane.

independent ILDs matched to the broadband RMS ILD of each HRTF; *ILD+ITD*, frequency-independent ILDs plus onset-threshold-estimated ITDs matched to each HRTF; *HRTF*, individually measured HRTFs; and $HRTF_{2.0}$, $HRTF_{3.0}$, and $HRTF_{4.0}$, augmented HRTFs with $\alpha = 2.0$, 3.0, and 4.0, respectively.

All sources were restricted to the horizontal plane. Target instruments were evenly divided across four categories (guitar, piano, synthesiser, and drums). Two geometries were tested: target at 0° azimuth with maskers divided between ±60°, and target at ±60° with one masker at 0° and the remaining maskers at the opposite 60° location. Eight samples were used per target category, each rendered under the seven rendering conditions and both spatial geometries. This was repeated 25 times, with each repeat using a unique HRTF from the SONICOM HRTF dataset corresponding to one individual listener [19]. Sample selection was randomised independently for each listener profile. For all stimuli, the target-to-masker level ratio was fixed at −6 dB RMS.

2.2.2. Auditory Model Analysis

To predict spatial release from masking, we employed two models available in the Auditory Modeling Toolbox [22]. First, SRM was estimated using the *vicente2020* binaural speech intelligibility model [23]. Second, we applied the *bischof2023* model, originally designed to predict the unmasking of reverberant complex tones in noise [24]. Both *vicente2020* and *bischof2023* combine better-ear SNR (BE SNR) and binaural unmasking (BU) to estimate effective SNR or total unmasking, respectively. Unlike *vicente2020*, *bischof2023* does not incorporate speech intelligibility weighting across frequency bands, providing a comparison that isolates the effect of speech weighting. It also utilises 12 ms analysis windows that allow it to implicitly account for the short-term monaural glimpsing effects explicitly modelled by *vicente2020*, making it better suited to non-stationary music signals than SRM models that assume relatively static inputs. SRM was defined as the dB benefit of spatial cues relative to the diotic condition, computed by subtracting the diotic score from the corresponding spatial condition score for each sample. Results were compared across conditions and target instrument categories.

2.2.3. Hearing Loss Simulation

Analysis with *vicente2020* was conducted using three hearing profiles: an N0 audiogram representing normal hearing, an N3 audiogram representing a moderate sensorineural hearing loss profile, and an N3 audiogram with samples pre-processed through a wide dynamic range compression (WDRC) hearing aid simulator. Hearing loss was modelled with *vicente2020*, which uses internal noise to capture threshold elevation and level-dependent coding deficits. Target and masking instruments were scaled together so that the broadband level of the maskers was fixed at 65 dB SPL prior to any amplification.

3. Results

3.1. SRM with Normal Hearing

Fig. 2 displays the mean SRM benefit predicted by *vicente2020* with an N0 audiogram, averaged across target instrument categories. All conditions elicit an SRM benefit relative to the diotic baseline. With a frontal target, individual HRTFs were predicted to provide 6.97 dB SRM, only a marginal increase of 0.18 dB over the ILD condition. Significantly, predicted SRM increased with the magnitude of HRTF augmentation: $HRTF_{4.0}$ showed the largest predicted SRM of 10.31 dB for a frontal target. Across conditions, SRM was higher when the target was located at ±60° than at 0°, on average by 1.88 dB.

For the augmented HRTF conditions, increases in SRM were mainly driven by improvements in BE SNR, with only minor reductions in BU. From individual HRTFs to $HRTF_{4.0}$, mean N0 SRM across target locations increased by 3.08 dB, comprising a 3.35 dB increase in BE SNR and a 0.27 dB reduction in BU. A similar pattern was observed under *bischof2023*, except that the ILD+ITD condition exceeded $HRTF_{3.0}$ in terms of SRM. Relative to the ILD condition, adding ITD increased SRM by 59% under *bischof2023* versus 29% under *vicente2020*.

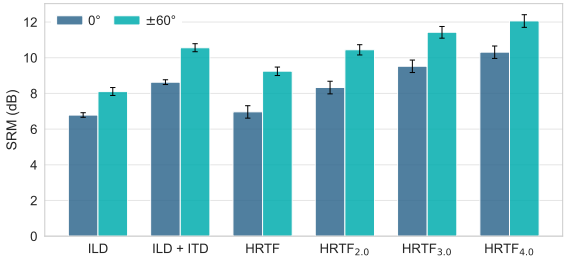

Figure 2: Mean predicted SRM relative to diotic baseline, split by target azimuth. Higher SRM indicates increased potential to discern the target instrument. Results are from *vicente2020* model with an N0 profile; error bars indicate bootstrapped 95% confidence intervals from 10,000 samples clustered by listener profile.

3.2. SRM with Hearing Loss

As indicated in Fig. 3, modelling an N3 loss with *vicente2020* resulted in a fall in mean SRM across conditions from 9.37 to 7.60 dB. Fig. 4 divides the predicted SRM into BE SNR and BU contributions, each relative to the diotic condition. This reveals, across conditions, an average 0.65 dB reduction in access to BU with an unaided N3 loss. Furthermore, BE SNR gains provided by HRTF augmentation were diminished: between individual HRTFs and $HRTF_{4.0}$, mean BE SNR gains fell from 3.35 dB to 1.49 dB for N0 and unaided N3 respectively, as access to higher frequency ILDs was lost. Though a significant SRM benefit remained for $HRTF_{4.0}$ versus the individual HRTFs with an N3 loss ($p < 0.001$), the magnitude of this gain decreased from 3.08 dB to 1.05 dB.

Pre-processing samples through WDRC to simulate a hearing aid partially restored audibility for the N3 profile, raising *vicente2020* mean effective SNR across conditions from −6.84 to −2.76 dB, but did not have a systematic effect on SRM across conditions. The SRM gain between $HRTF_{4.0}$ and individual HRTFs remained unchanged from the unaided N3 result ($\Delta = +0.01$ dB, $p = 0.81$). On average across all dichotic mixes, WDRC processing also resulted in 2.87 dB distortion to short-term (24 ms) broadband ILD. Larger ILD distortion correlated with reductions in SRM relative to N0 ($r = 0.31$, $p < 0.001$).

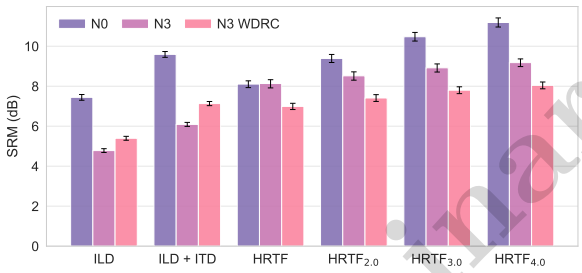

Figure 3: Mean SRM predicted by *vicente2020* averaged over target locations, split by hearing profile: normal hearing (N0), unaided hearing loss (N3), aided hearing loss (N3 WDRC). Error bars show bootstrapped 95% confidence intervals.

3.3. Effect of Target Instrument

Under N0, the SRM gain provided by $HRTF_{4.0}$ relative to individual HRTFs differed across target instruments (Friedman $p = 0.003$), with the synthesiser targets showing significantly higher SRM gain than drum targets (3.39 dB versus 2.80 dB, Holm-adjusted $p = 0.007$). In the N3 and N3 aided conditions, no significant differences emerged between instruments.

4. Discussion

4.1. Spatial Cue Contributions to MSA

A data-driven augmentation method was outlined, in which principal components derived from a population of measured HRTFs were scaled to enhance frequency-dependent ILDs in individual HRTFs. For an N0 listener, auditory models predicted greater SRM with augmented than with individual HRTFs, suggesting that augmentation may support clearer instrument separation than is achievable using natural spatial cues. While augmentation benefits were largely attributed to elevated BE SNR, SRM predictions for an N0 listener also increased with the introduction of ITDs: predicted SRM was greater under the ILD+ITD condition versus the ILD condition. This deviates from the behavioural results of [25], who reported that ITDs added marginal SRM benefit in a speech-on-speech task when ILDs were already present. This discrepancy may be tied to high informational masking creating performance limits in that study, which the present approach does not account for. However, the significance of ITD-based equalisation-cancellation mechanisms in music remains an open question; unlike BU speech models designed to cancel competing talkers, MSA may prioritise the ability to hear multiple instruments simultaneously.

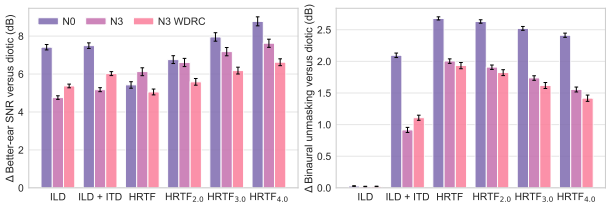

Figure 4: Mean change in BE SNR (left) and BU (right) relative to diotic condition, under *vicente2020*. Results averaged over target locations and split by simulated hearing profile. Error bars show bootstrapped 95% confidence intervals.

Across results, predicted SRM was influenced by the frequency distribution of spatial cues. For instance, *bischof2023* predicted a larger relative benefit for the ILD+ITD condition than *vicente2020*, plausibly because speech-weighting curves down-weight low-frequency SRM. Since music has more variable long-term spectra than speech [26], low frequencies may contribute more to SRM than speech-weighting functions assume, making a corresponding behavioural result likely. Additionally, for N0 listeners, *vicente2020* predicted comparable SRM for the ILD and HRTF conditions, even though HRTFs also convey ITDs. This arises as the ILD condition was matched to the broadband (80Hz-16kHz) ILD of each individual HRTF, whereas ILDs in a HRTF vary with frequency. As such, target instrument spectral centroids fell at frequencies where ILDs in the HRTF condition were smaller than those in the flat ILD condition, on average by 5.09 dB at ±60°. Together, these results suggest that spatial cue benefits for MSA rely on how cue magnitude aligns with the target spectrum.

4.2. Audibility Limits SRM Benefits

With an unaided N3 loss, the SRM benefit of HRTF augmentation persisted but was largely diminished

as elevated internal noise rendered higher frequency ILDs inaudible. Individual HRTFs were also predicted to yield greater SRM than broadband-matched ILD+ITD cues, reversing the N0 ordering. Further analysis suggested this was driven by an audibility effect. Under N3, elevated internal noise around 2-4 kHz shifted the primary determinant of SRM from head-shadowing advantages to pure target audibility. HRTF spectral colouration increased target energy in this region, allowing brief glimpses above the internal noise floor, whereas ILD+ITD provided no comparable spectrally local improvement. Hearing aid processing partially restored audibility in this region, thereby reinstating the N0 condition ordering. Overall though, average SRM with WDRC remained below normal-hearing levels, likely due to residual hearing loss effects and the interaural cue distortions introduced by unlinked bilateral WDRC processing.

4.3. Limitations

There are notable limitations to the presented modelling approach. First, SRM in music listening may be modulated by additional auditory scene analysis cues such as temporal modulation, harmonic structure, and pitch differences, in ways not captured by models of speech and complex tone unmasking. The hearing loss modelling is also approximate, and does not explicitly account for auditory filter broadening and reduced temporal fine structure processing [23]. These omissions may fail to capture cues that are relevant for SRM in music listening. In particular, the lack of substantial differences in SRM gain between target instruments observed here, especially under an N3 audiogram, seems unlikely to be replicated behaviourally. Enhanced low-frequency ILDs have improved localisation in listeners with hearing loss only for stimuli with distinct envelope fluctuations [27], suggesting that instrument-dependent temporal-envelope characteristics could influence behavioural results.

Second, objective models neglect higher-level processes. Given evidence that the auditory scene analysis of music and speech share perceptual mechanisms [28], behavioural SRM may diverge from model predictions as structural cues and listener adaptation modulate informational masking. For instance, [29] observed significantly greater median-plane SRM with individual HRTFs than mannequin HRTFs. [30] found no speech intelligibility benefit from interaural magnification, partly attributing this to distorted spatial cues introduced by the magnification algorithm. Such unfamiliar cues may be harder to process, particularly given potential relationships between localisation accuracy and SRM performance [31]. Augmented HRTFs may similarly depart from a listener's learned spatial mapping, potentially explaining why [13] observed smaller behavioural SRM gains than predicted by auditory models. Nevertheless, since HRTF augmentation improved speech intelligibility in that study, we expect the present approach to yield benefits in behavioural music scene analysis tests.

5. Conclusion

A framework was introduced by which spectral cues embedded in HRTFs were augmented to increase the spatial release from masking of target instruments in music. Auditory model predictions indicated a benefit to the approach, with music rendered using augmented HRTFs achieving elevated SRM compared with individual HRTFs. This advantage persisted under simulation of moderate sensorineural hearing loss but was substantially reduced. Hearing aid processing partially restored overall audibility, but did not meaningfully recover augmentation-related SRM gains.

6. Acknowledgments

This work was supported by an EPSRC CASE conversion PhD scholarship, co-funded by Sonova. Our thanks to Michel Bürgel and Kai Siedenburg for their provision of the stimuli from [7].

References

[1] A. E. Greasley et al., "Using hearing aids for music: A UK survey of challenges and strategies," *Trends in Hearing*, vol. 30, pp. 1–33, 2026.

[2] A. Greasley, H. Crook, and R. Fulford, "Music listening and hearing aids: Perspectives from audiologists and their patients," *International Journal of Audiology*, vol. 59, no. 9, pp. 694–706, 2020.

[3] A. S. Bregman, *Auditory Scene Analysis: The Perceptual Organization of Sound*. MIT press, 1994.

[4] J. F. Culling and M. A. Stone, "Energetic masking and masking release," in *The Auditory System at the Cocktail Party*, Springer, 2017, pp. 41–73.

[5] G. Kidd Jr and H. S. Colburn, "Informational masking in speech recognition," in *The Auditory System at the Cocktail Party*, Springer, 2017, pp. 75–109.

[6] A. Ihlefeld and B. Shinn-Cunningham, "Spatial release from energetic and informational masking in a selective speech identification task," *The Journal of the Acoustical Society of America*, vol. 123, no. 6, pp. 4369–4379, 2008.

[7] R. Hake, M. Bürgel, N. K. Nguyen, A. Greasley, D. Müllensiefen, and K. Siedenburg, "Development of an adaptive test of musical scene analysis abilities for normal-hearing and hearing-impaired listeners," *Behavior Research Methods*, vol. 56, no. 6, pp. 5456–5481, 2024.

[8] R. Hake et al., "Perception of recorded music with hearing aids: Compression differentially affects musical scene analysis and musical sound quality," *Trends in Hearing*, vol. 29, pp. 1–21, 2025.

[9] A. J. Benjamin and K. Siedenburg, "Effects of spectral manipulations of music mixes on musical scene analysis abilities of hearing-impaired listeners," *PLOS One*, vol. 20, no. 1, e0316442, 2025.

[10] N. Durlach and X. Pang, "Interaural magnification," *The Journal of the Acoustical Society of America*, vol. 80, no. 6, pp. 1849–1850, 1986.

[11] B. Kollmeier and J. Peissig, "Speech intelligibility enhancement by interaural magnification," *Acta Oto-Laryngologica*, vol. 109, no. sup469, pp. 215–223, 1990.

[12] B. N. Richardson, J. M. Kainerstorfer, B. G. Shinn-Cunningham, and C. A. Brown, "Magnified interaural level differences enhance binaural unmasking in bilateral cochlear implant users," *The Journal of the Acoustical Society of America*, vol. 157, no. 2, pp. 1045–1056, 2025.

[13] N. Marggraf-Turley, N. H. Pontoppidan, and L. Picinali, "The impact of individual head-related transfer function augmentation on spatial release from masking," *Hearing Research*, vol. 466, 2025, Art. no. 109402.

[14] D. S. Brungart and G. D. Romigh, "Spectral hrtf enhancement for improved vertical-polar auditory localization," in *2009 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, IEEE, 2009, pp. 305–308.

[15] T. L. Arbogast, C. R. Mason, and G. Kidd Jr, "The effect of spatial separation on informational masking of speech in normal-hearing and hearing-impaired listeners," *The Journal of the Acoustical Society of America*, vol. 117, no. 4, pp. 2169–2180, 2005.

[16] T. Goossens, C. Vercammen, J. Wouters, and A. van Wieringen, "Masked speech perception across the adult lifespan: Impact of age and hearing impairment," *Hearing Research*, vol. 344, pp. 109–124, 2017.

[17] K. Siedenburg, S. Röttges, K. C. Wagener, and V. Hohmann, "Can you hear out the melody? Testing musical scene perception in young normal-hearing and older hearing-impaired listeners," *Trends in Hearing*, vol. 24, pp. 1–15, 2020.

[18] D. J. Kistler and F. L. Wightman, "A model of head-related transfer functions based on principal components analysis and minimum-phase reconstruction," *The Journal of the Acoustical Society of America*, vol. 91, no. 3, pp. 1637–1647, 1992.

[19] K. C. Poole et al., "The extended sonicom hrtf dataset and spatial audio metrics toolbox," in *Proc. 11th Conv. EAA Forum Acusticum*, 2025, pp. 6483–6486.

[20] R. Eastgate, L. Picinali, H. Patel, and M. D'Cruz, "3d games for tuning and learning about hearing aids," *The Hearing Journal*, vol. 69, no. 4, pp. 30–32, 2016.

[21] R. Bittner, J. Wilkins, H. Yip, and J. P. Bello, *Medleydb 2.0 audio*, Zenodo, Aug. 2016.

[22] P. Majdak, C. Hollomey, and R. Baumgartner, "Amt 1.6: A toolbox for reproducible research in auditory modeling," *Acta Acustica*, vol. 6, p. 19, 2022.

[23] T. Vicente, M. Lavandier, and J. M. Buchholz, "A binaural model implementing an internal noise to predict the effect of hearing impairment on speech intelligibility in non-stationary noises," *The Journal of the Acoustical Society of America*, vol. 148, no. 5, pp. 3305–3317, 2020.

[24] N. F. Bischof, P. G. Aublin, and B. U. Seeber, "Fast processing models effects of reflections on binaural unmasking," *Acta Acustica*, vol. 7, p. 11, 2023.

[25] H. Glyde, J. M. Buchholz, H. Dillon, S. Cameron, and L. Hickson, "The importance of interaural time differences and level differences in spatial release from masking," *The Journal of the Acoustical Society of America*, vol. 134, no. 2, EL147–EL152, 2013.

[26] M. Chasin and F. A. Russo, "Hearing aids and music," *Trends in Amplification*, vol. 8, no. 2, pp. 35–47, 2004.

[27] B. C. Moore, A. Kolarik, M. A. Stone, and Y.-W. Lee, "Evaluation of a method for enhancing interaural level differences at low frequencies," *The Journal of the Acoustical Society of America*, vol. 140, no. 4, pp. 2817–2828, 2016.

[28] R. Hake, D. Müllensiefen, and K. Siedenburg, "Individual differences in auditory scene analysis abilities in music and speech," *Scientific Reports*, vol. 15, no. 1, 2025, Art. no. 24048.

[29] D. González-Toledo, M. Cuevas-Rodríguez, T. Vicente, L. Picinali, L. Molina-Tanco, and A. Reyes-Lecuona, "Spatial release from masking in the median plane with non-native speakers using individual and mannequin head related transfer functions," *The Journal of the Acoustical Society of America*, vol. 155, no. 1, pp. 284–293, 2024.

[30] T. de Taillez, G. Grimm, B. Kollmeier, and T. Neher, "Acoustic and perceptual effects of magnifying interaural difference cues in a simulated 'binaural' hearing aid," *International Journal of Audiology*, vol. 57, no. sup3, S81–S91, 2018.

[31] N. K. Srinivasan, K. M. Jakien, and F. J. Gallun, "Release from masking for small spatial separations: Effects of age and hearing loss," *The Journal of the Acoustical Society of America*, vol. 140, no. 1, EL73–EL78, 2016.