Hearing Research 466 (2025) 109402

Contents lists available at ScienceDirect

Hearing Research

journal homepage: www.elsevier.com/locate/heares

Research paper

The impact of individual head-related transfer function augmentation on spatial release from masking

Nils Marggraf-Turley [a],*, Niels H. Pontoppidan [b], Lorenzo Picinali [a]

[a] Dyson School of Engineering, Imperial College London, South Kensington Campus, London, SW7 2AZ, United Kingdom
[b] Eriksholm Research Centre, Oticon A/S, Snekkersten, 3070, Denmark

ARTICLE INFO

Dataset link: 10.5281/zenodo.15198918

Keywords:
Head-related transfer function
Spatial release from masking
Interaural cue magnification
Binaural hearing

ABSTRACT

Binaural hearing underpins effective communication in complex acoustic environments by increasing listeners' abilities to segregate concurrent sound sources. In certain conditions, interaural magnification of binaural cues has been shown to improve speech intelligibility in competing target masker scenarios, yet existing methods primarily comprise hearing aid algorithms, which, due to processing constraints, cause unwanted artefacts. Moreover, the perceptual effects of applying interaural magnification directly to a person's own head-related transfer function (HRTF) remain unclear.

We revisit a previous method for magnifying interaural cues, apply it to individual HRTFs, and propose a modified version designed to reduce spatial distortions. We evaluate both approaches using auditory models and listening tests involving speech-on-speech masking with a midline target and symmetrical maskers. Behavioural results show that although magnifying interaural phase and level differences improves overall speech intelligibility compared to unmodified HRTFs, it does not significantly enhance spatial release from masking (SRM). Meanwhile, auditory model predictions suggest greater SRM gains than those observed behaviourally, suggesting an influence of listener familiarity with their individual spatial cues. Finally, interaural magnification benefits diminish for listeners who already exhibit strong SRM with their own HRTFs.

1. Introduction

Spatial auditory perception relies on interaural differences in time (ITD) and level (ILD), combined with monaural spectral cues, to facilitate sound source localisation and the segregation of concurrent auditory streams. This segregation improves speech intelligibility (SI) in cocktail-party contexts (Cherry, 1953). A key mechanism behind this is spatial release from masking (SRM), which exploits better-ear signal-to-noise ratio (SNR) and binaural unmasking (BU) driven by interaural differences (Durlach, 1963).

The head-related transfer function (HRTF) encapsulates these interaural and monaural spectral cues, thereby describing each listener's morphological filtering. Consequently, HRTFs are widely employed for simulating virtual auditory environments. Despite being considered essential for accurate spatialisation, individual HRTFs do not always yield optimal perceptual results (Armstrong et al., 2018), and practical applications frequently rely on non-individual HRTFs due to the complexity of direct measurement.

1.1. The influence of HRTF on SRM

The extent to which an individual benefits from SRM appears to depend on their HRTF. Ahrens et al. (2021) explored this phenomenon using a computational auditory model (Jelfs et al., 2011) to predict SRM across HRTFs from the LISTEN and CIPIC databases, revealing inter-HRTF variability of up to 4–5.9 dB. This indicates that certain HRTFs may provide more favourable acoustic cues for speech segregation than others. Moreover, their results showed that the widely used KEMAR HRTF provided higher SRM than the median of individual HRTFs, reinforcing the notion that certain non-individual HRTFs might outperform individual ones in specific applications.

While these findings support using HRTFs that maximise SRM for simulating complex environments, evidence also indicates that spatial cue disparities alone do not fully explain SRM differences. The findings of Cuevas-Rodriguez et al. (2021) and Orduña-Bustamante et al. (2018) suggest that listener familiarity with spatial cues may also play a role. Cuevas-Rodriguez et al. (2021) conducted a speech-in-noise study in a virtual cocktail party scenario, with a frontal speech target and

* Corresponding author.
E-mail addresses: n.marggraf-turley23@imperial.ac.uk (N. Marggraf-Turley), npon@eriksholm.com (N.H. Pontoppidan), l.picinali@imperial.ac.uk (L. Picinali).

https://doi.org/10.1016/j.heares.2025.109402
Received 8 May 2025; Received in revised form 11 August 2025; Accepted 11 August 2025
Available online 20 August 2025

lateral maskers at ±90°. Participants used measured non-individual and synthetic HRTFs (modelled with ILD and ITD cues). Their findings revealed significant inter-individual variability in speech reception thresholds across different HRTFs, with some subjects experiencing measurable SRM benefits from certain HRTFs, while others showed reduced performance. Crucially, no single HRTF emerged as optimal across all listeners in contrast to the model-based results of Ahrens et al. (2021). Similarly, Orduña-Bustamante et al. (2018) examined SI differences between individual and non-individual HRTFs in noise and reverberation. They found that non-individual HRTFs led to moderate but significant reductions in intelligibility (up to 12.1% with interaurally correlated noise). They used Complex Spectral Distortion (CSD), which computes the root-mean-square (RMS) difference between the complex frequency response of the target and reference spectra, thereby accounting for phase and magnitude differences. However, they observed that even after accounting for CSD, individual HRTFs still provided an intelligibility advantage, suggesting that long-term adaptation to an individual's spatial cues may contribute to speech perception in noise. Complementary findings were observed for median plane SRM by González-Toledo et al. (2024). They investigated SRM in the median plane, where binaural cues are minimal, and monaural spectral cues dominate. Their study tested SRM performance for individual versus KEMAR HRTFs in a speech-on-speech task with non-native English speakers. They found that SRM was significantly higher when using individual HRTFs compared to KEMAR HRTFs, suggesting that listeners could take better advantage of their own spectral cues. Additionally, this effect was stronger in participants with lower English proficiency, suggesting a potential role of spatial cue familiarity in compensating for linguistic processing difficulties. Similar to the earlier work by Ahrens et al. (2021), these authors also used the Jelfs et al. (2011) computational model of SRM alongside spectral distortion to predict the expected SRM. However, neither metric fully accounted for the observed results, with certain individuals performing better with individual HRTFs even though the Jelfs et al. (2011) SRM and spectral distortion were higher with the non-individual HRTF, which indicates familiarity with their own cues outweighed the magnitude of spatial cues.

1.2. HRTF augmentation

It appears that both spatial cue magnitude and familiarity contribute to SRM, raising a pertinent question: Can an individual's HRTF be modified to enhance SRM beyond their natural capabilities? While many studies focus on approximating an individual HRTF (Miccini and Spagnol, 2020; Kaneko et al., 2016), relatively few have explored HRTF augmentation.

Various augmentation strategies have been proposed, including spectral enhancement to increase direction-dependent cues for improved vertical localisation (Brungart and Romigh, 2009), ITD magnification in virtual reality environments (Guiraud et al., 2024), frequency-selective amplification to reduce front–back confusions (So et al., 2011), perceptually guided HRTF adjustment incorporating audio quality metrics (Yao and Chen, 2013), and frequency scaling to align individual and non-individual HRTFs (Middlebrooks, 1999). However, these studies primarily target localisation accuracy, with limited focus on potential benefits for SRM.

As Ahrens et al. (2021) noted, HRTFs optimised for complex cocktail-party-like auditory scenes could be beneficial. However, most augmentation studies focus exclusively on localisation accuracy rather than SRM enhancement. To the best of the authors' knowledge, the only study considering HRTF augmentation for SRM enhancement is the work of Pirhosseinloo and Kokkinakis (2016). Their interaural magnification (IM) algorithm magnifies interaural differences between left and right HRTF spectra, similar to the algorithm of Durlach and Pang (1986) but instead modifying HRTFs before rendering spatialised audio rather than applying processing to a binaural signal post hoc. However, they did not present perceptual evaluations, leaving open questions about its effectiveness for SRM.

1.3. Interaural magnification

In the context of hearing aid (HA) processing, IM has been demonstrated to be somewhat effective for enhancing interaural disparities to improve horizontal plane SRM. Durlach and Pang (1986) introduced a signal-processing scheme that exponentially magnifies IPD and ILD cues, a concept later evaluated and implemented by Kollmeier and Peissig (1990) who found that IM improved SI in some normal-hearing (NH) listeners. De Taillez et al. (2018) expanded on this research by implementing IM within a simulated binaural HA framework. Their study evaluated SI, spatial perception, and signal quality in elderly normal-hearing and hearing-impaired (HI) listeners. While IM successfully increased the angular separation of competing sources, it did not yield improvements in intelligibility. They noted that the exponential ILD+IPD magnification distorted spatial cues by spreading them, which was accentuated in diffuse noise environments. They suggested that this reduced any potential benefit from IM for speech perception. Other studies have explored related approaches for cochlear implant users, with ILD enhancement showing intelligibility benefits in some cases (Brown, 2014; Lopez-Poveda et al., 2016; Richardson et al., 2025).

However, these studies primarily focused on HI listeners, who lack access to their individual HRTFs due to HA processing (Brown, 2014) or integrate non-individual HRTFs into their algorithm (Lopez-Poveda et al., 2016). Conversely, the studies which consider NH listeners do not compare with their individual HRTF (Durlach and Pang, 1986; Kollmeier and Peissig, 1990; De Taillez et al., 2018) and therefore IM benefits observed in non-individual HRTF conditions may not directly translate to individual listening, since individualised spatial cue familiarity may be a factor.

1.4. Summary and aims of the present study

From the review outlined in the previous sections, it is evident that SRM varies with HRTF type, and the magnitude and, perhaps, the familiarity of spatial cues seem to play an important role. Existing HRTF augmentation approaches primarily target localisation accuracy, and those that consider SRM often lack direct perceptual evaluations. Conversely, some HA-based IM techniques have demonstrated effectiveness in improving SI; however, these studies have primarily been conducted in contexts where individual HRTFs are either bypassed (as in HA users) or not explicitly considered (as in non-individual HRTF processing). Consequently, it remains unclear whether the benefits of IM generalise to individual HRTF modifications and whether these augmentations can improve SRM beyond a listener's natural HRTF. To address this gap, we present an exploratory study investigating whether IM applied directly to an individual HRTF can enhance SRM beyond individual performance. We implement (De Taillez et al., 2018)'s version of the hearing-aid-based IPD and ILD magnification approach of Durlach and Pang (1986) but apply it directly to the HRTF rather than a binaural signal comprising already mixed sources. Additionally, since the window-based binaural stream processing constraints typical in HA algorithm design do not apply when augmenting HRTFs directly, we propose a modified IM algorithm incorporating ITD+ILD magnification to mitigate spatial cue spreading as reported by De Taillez et al. (2018). A speech-on-speech, symmetrical masker SRM listening test is conducted with 11 participants under three spatial conditions: (1) individual HRTF, (2) individual HRTF with Durlach and Pang (1986) IM, and (3) individual HRTF with our modified IM approach. Observed SRM was compared to predictions from auditory models and numerical metrics to assess whether factors beyond interaural cue magnitude influence SRM. Our hypotheses are as follows, with outcomes summarised in Table 1:

- H1 IM of an individual HRTF enhances horizontal plane SRM beyond individual performance.

Table 1
Summary of hypotheses outcomes.

Hyp.	Summary outcome	Supported
H1	IM increased SI for C2 but not SRM due to increased SI under co-located conditions.	No
H2	Auditory model predictions showed reduced spatial spreading and higher SRM/spread ratio for C3 (ITD+ILD) than C2 (IPD+ILD), although likely below perceptual thresholds.	Partial
H3	Numerical metrics failed to predict SRM gain.	Yes

H2 ITD+ILD-based IM will cause less spatial cue spreading than IPD+ILD.
H3 Numerical signal differences will not completely account for SRM differences between individual and augmented conditions.

2. Materials and methods

2.1. Interaural magnification

Individual *FreeFieldCompMinPhase* HRTFs from the SONICOM database (Engel et al., 2023) at 48 kHz were used for each participant and were the baseline for IM. Two magnified spatial conditions were employed to test hypothesis H1:

C2 The IM algorithm of Durlach and Pang (1986) with a crossover frequency $\omega_c = 1.5$ kHz to minimise phase ambiguities following De Taillez et al. (2018). IPD magnification is applied for $\omega \leq \omega_c$ and ILD magnification for $\omega > \omega_c$.
C3 ITD magnification applied to the contralateral ear instead of IPD, while still applying ILD magnification for $\omega > \omega_c$.

The mathematical formulations for these conditions are provided in Appendix.

2.2. Participants

Participants (n = 11) with self-reported normal hearing (eight male, three female) took part in the study. All participants, except two (P05 and P06), reported native or fluent proficiency in English; P05 and P06 described their proficiency as intermediate.

2.3. Speech stimuli

The experiment employed a speech-on-speech SRM test with symmetrical masker configurations. The target was presented using HRTFs corresponding to 0° azimuth and 0° elevation, while masker stimuli were positioned on the horizontal plane at azimuths of ±0°, ±5°, ±10°, ±15°, and ±20°.

Speech stimuli were drawn from the Coordinate Response Measure (CRM) corpus (Bolia et al., 2000), which comprises sentences structured as:

"Ready [call sign], go to [colour] [number] now".

The corpus contains eight call signs, four colours, and eight numbers, with recordings from four male and four female speakers. The target speech always used the call sign *Charlie*. Behavioural studies show that listeners can exploit spectral detail beyond 8 kHz for lateral localisation (Morikawa and Hirahara, 2010; Brungart and Simpson, 2009). Consequently, the original CRM recordings (low-pass filtered at 18 kHz) were used to preserve the broadband characteristics required for IM, rather than the standard version filtered at 8 kHz. For each trial, target and masker stimuli were selected such that they were spoken by different speakers of the same gender, and neither masker shared the target's call sign.

For the individual HRTFs, an additional ITD-matching condition was incorporated. Specifically, after applying IM for condition C3, the final ITD was computed, and the unmodified individual HRIR in the frontal hemisphere with the closest ITD alignment was selected. This condition was included to examine the effects of ILD magnification in isolation. All stimuli were convolved with individual headphone equalisation filters from the SONICOM database, and head tracking was not employed, thereby ensuring a fixed sound scene.

2.4. Behavioural experiment procedure

Stimuli presentation was controlled by a custom Python script that displayed a GUI for recording participants' responses (see Fig. 1). Participants were instructed to identify the colour and number spoken by the target voice, which was always presented directly in front of them. To familiarise participants with the task, a familiarisation phase was conducted prior to the main experiment, similar to Daugintis et al. (2024). The initial 10 training trials featured only the target speech, while subsequent trials introduced a single masker alongside the target, with masker locations randomly sampled without replacement from the set of locations under test. Immediate feedback on response accuracy was provided after each training trial. The main experiment was structured into 20 blocks of 24 trials each, with different spatial conditions and masker locations presented in random order such that each masker location was tested 30 times per spatial condition. The experiment lasted approximately 40 min, and participants were informed that they could take breaks between stimuli and blocks to minimise fatigue.

2.5. Numerical analysis

Following Ahrens et al. (2021), we employed the revision of the SRM prediction model by Lavandier and Culling (2010) proposed by Jelfs et al. (2011). This model predicts SRM by combining the benefits of better–ear SNR and BU. Since the model operates directly on the HRIR, it allowed us to evaluate the effect of IM on the SNR and BU components of SRM, without considering target and masker stimuli variations. Total SRM was computed along with its BU and SNR components for both a symmetrical and single-masker configuration, with the target always at 0° azimuth and elevation, as in our behavioural experiment. Masker azimuths from 0° to 45° in intervals of 5° were tested. Only 0° to 20° was tested behaviourally due to experiment length constraints.

De Taillez et al. (2018) demonstrated that the IM approach proposed by Durlach and Pang (1986) disperses direction-of-arrival (DoA) estimates as derived from the auditory model of Dietz et al. (2011), which we refer to here as spatial spreading. Consequently, to compare spatial spreading between the spatial conditions in the study, thereby investigating H2, we adopted a similar analysis.

The model of Dietz et al. (2011) estimates the DoA of sound sources through a series of peripheral auditory processing stages, including middle-ear filtering, basilar membrane band-pass filtering, cochlear compression, and mechano-electrical transduction. It extracts ITDs and ILDs across frequency channels, maps ITDs to azimuth angles, and resolves ambiguities using ILDs. A binaural coherence threshold is applied so that only highly coherent signal segments (>0.98) contribute to the DoA estimation. ITD estimates are mapped to azimuth angles using a lookup table calibrated to each listener's unmodified individual HRTF. The azimuth estimates obtained over the full duration of each stimulus were then aggregated into histograms to quantify the frequency of estimated source locations. The spatial spread of estimated azimuths was assessed by computing the standard deviation of the DoA distributions separately for the left (−90° to 0°) and right (0° to 90°) quadrants. The target was removed to enable separate analyses of the spatial spreading in each quadrant without confounding with DoA of

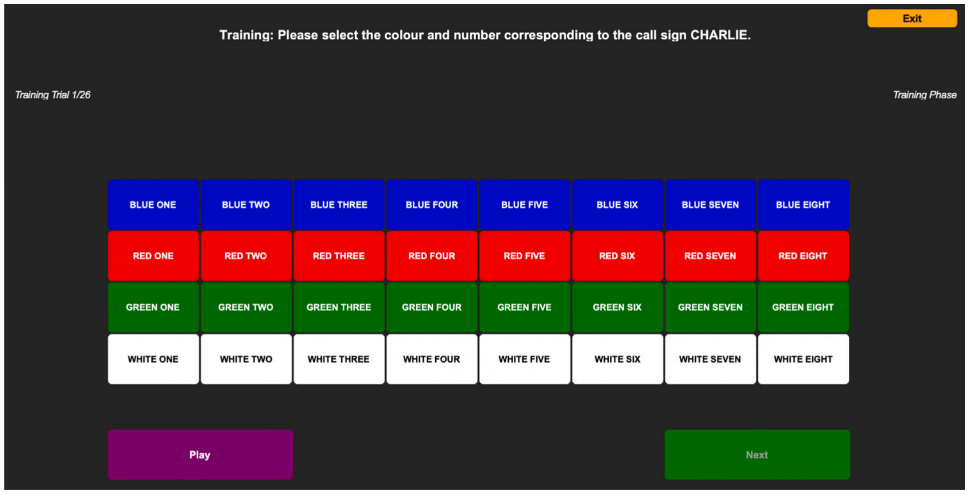

Fig. 1. Graphical user interface to select the colour-number combination of the target sentence.

the target at azimuth 0°. The average spread across both quadrants was taken.

To investigate H3, three numerical measures were employed to relate physical signal differences to behavioural SRM. The first measure was the SRM prediction model of Lavandier and Culling (2010) via the *lavandier2022* function in the Auditory Modelling Toolbox (Majdak et al., 2022). This model applies the same processing as Jelfs et al. (2011), as outlined above, but operates on the final binaural signals instead of the HRIRs as in Jelfs et al. (2011), since we wanted to account for the different target and masker spectra. We refer to this as LCJ SRM.

We also employed the Binaural Auditory-Model-Based Quality Prediction (BAM-Q) model of Flessner et al. (2017). In contrast to the model of Jelfs et al. (2011), which calculates BU and SNR components, Flessner et al. (2017) adopted an adapted version of the binaural feature extraction model proposed by Dietz et al. (2011). This model extracts ITD, ILD, and Interaural Vector Strength (IVS), a measure related to interaural correlation (IC). Flessner et al. (2017) then use a regression-based back-end to predict binaural quality degradation based on subjective ratings. Although this model was not explicitly evaluated within a strict SI paradigm, the model correlated highly ($r = 0.92$) with listeners' ratings of the overall difference between stimuli processed with the IM approach of Kollmeier and Peissig (1990). For each subject and spatial condition, BAM-Q was calculated between target and masker stimuli and averaged across all trials. The value obtained under co-located target-masker configurations was subtracted from each separated conditions to yield a predicted SRM measure.

Finally, we employed CSD as used by Orduña-Bustamante et al. (2018) as a purely numerical metric of HRTF differences, without modelling auditory system processing. CSD quantifies the discrepancy between the real and imaginary components of two signals by computing their RMS difference. Predicted SRM for each azimuth was calculated as the CSD between the masker and target HRTFs.

In addition to computing the predicted SRM at each masker azimuth, we also calculated the predicted SRM gain at each azimuth, defined as the difference between the SRM of conditions C2 and C3 and that of the individual HRTFs. By comparing SRM gains across the different models, we can further assess whether improvements in SRM from IM can be explained by looking only at numerical measures.

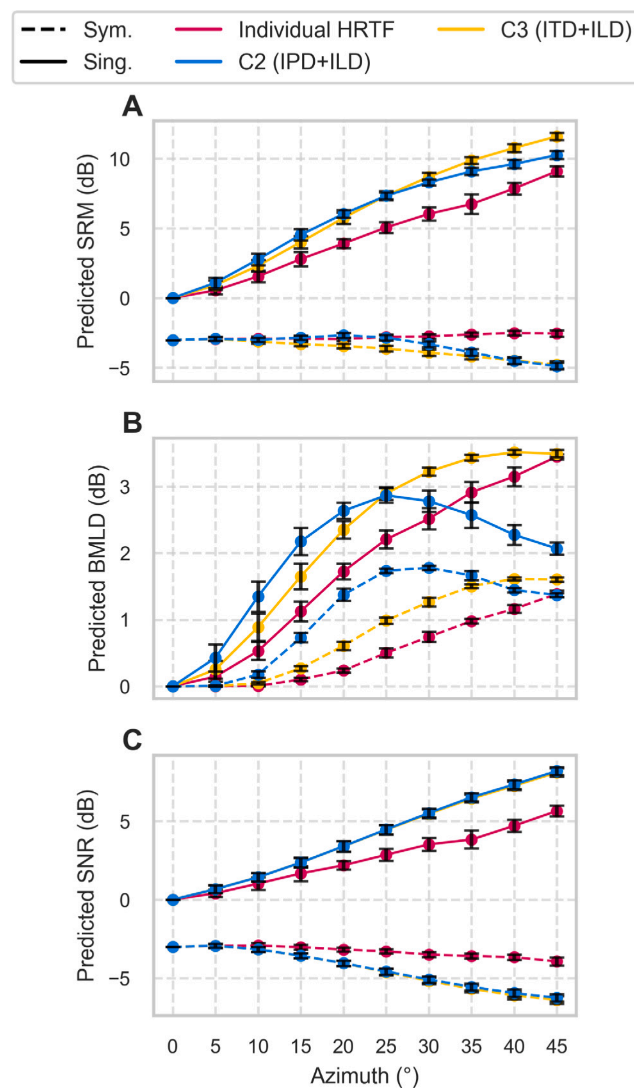

Fig. 2. Predicted SRM (A), binaural masking level difference (B) and SNR (C), for symmetrical (Sym.) and single (Sing.) masker configurations across different azimuths. SRM is plotted for Individual and IM conditions: C2 (IPD+ILD magnification) and C3 (ITD+ILD magnification). Error bars represent 95% CI across participants.

3. Results

3.1. Auditory model evaluation

Fig. 2(A) displays the predicted SRM, computed according to Jelfs et al. (2011), as a function of masker azimuth. Error bars represent

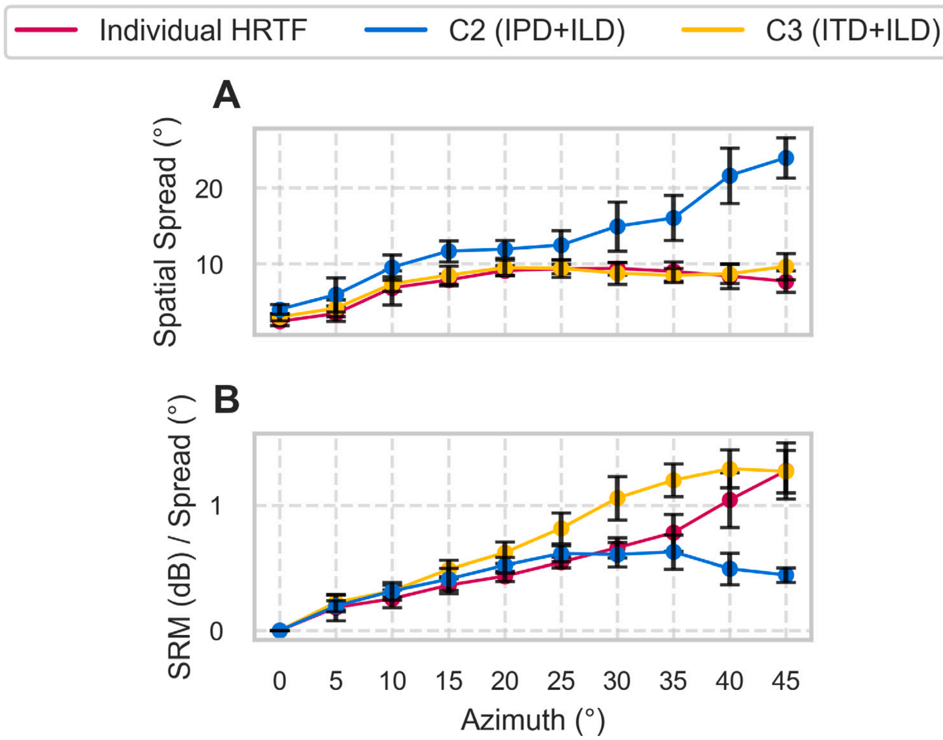

Fig. 3. Spatial spreading (standard deviation of direction-of-arrival estimates) across masker azimuths for individual and magnified HRTFs (C2 & C3). Error bars represent 95% CI across participants.

95% confidence intervals based on the 11 individual HRTFs used. Under the single-masker configuration, unmodified individual HRTFs consistently yield the lowest predicted SRM across all azimuths. In contrast, conditions C2 and C3 intersect at approximately 25°: C2 yields higher SRM than C3 at azimuths below 25°, whereas beyond 25°, C3 becomes dominant. For the symmetrical configuration, SRM remains within 0.5 dB for azimuths up to 20°, after which both C2 and C3 decline, reaching a minimum at 45° that is approximately 2.5 dB lower than for the individual HRTFs.

Fig. 2(B) isolates the BU contribution. In the single-masker configuration, C2's BU component initially exceeds that of the Individual HRTF by 0.5–1.25 dB; however, its advantage decreases relative to C3, resulting in a crossover at 25°. Beyond 35°, the BU for C2 falls below the Individual HRTF, reaching a minimum of 2 dB at 45°, at which point C2 and the Individual HRTF converge. A similar trend is observed for the symmetrical configuration, though C2's decline relative to C3 occurs at a more lateral azimuth (40°).

Fig. 2(C) depicts the SNR contribution to SRM. Under the single masker configuration, both C2 and C3 exhibit nearly identical SNR improvements that consistently exceed those of the Individual HRTF, reaching a maximum advantage of approximately 2.5 dB at 35° azimuth. The inverse is observed under the symmetrical masker configuration, with the Individual HRTF exhibiting an almost constant SNR of approximately −3 dB, while both C2 and C3 produced lower SNRs, with a minimum of 2.5 dB below Individual HRTFs at an azimuth of 45°.

Fig. 3(A) presents the predicted angular spread across azimuth for each condition, focusing on the single-masker scenario due to the limited SRM variation in the symmetrical configuration. C2 shows the largest predicted spread at all azimuths, peaking at approximately 23° when the masker is at 45°. In contrast, C3 and the Individual HRTF remain closely matched, differing by less than 1° across azimuths until 45°, where C3's median spread exceeds that of the individual HRTF by approximately 3°. Fig. 3(B) shows the ratio of predicted SRM to angular spread. Below 20° azimuth, the SRM gain per degree of spread is under 0.5 dB for all conditions. Beyond 20°, C3 demonstrates approximately 0.5–1 dB higher SRM per degree of spread relative to C2. By 45°, the Individual HRTF and C3 converge in terms of SRM-per-spread, whereas C2 remains approximately 1 dB lower.

3.2. Behavioural SRM and SI

Fig. 4 shows boxplots of behavioural SI and SRM, averaged across all trials per subject, as a function of masker azimuth. SRM was calculated as the difference between SI under spatially separated and co-located conditions. For most azimuths, the median SI was higher in C2 compared to C3 and individual conditions, except at 10° where C3's median SI was approximately 10% higher than individual. Notably, the individual condition matched to the ITD of C3 exhibited higher median SI than C3 at 20° (83% vs. 70%). Surprisingly, SI differences were also present at 0° azimuth, even though IM was applied equally to both target and masker HRTFs at that location.

SRM increased with masker azimuth across all conditions. At 5° azimuth, median SRM ranged between 3% and 18% for all conditions, with C2 showing the greatest interquartile range. As azimuth increased, inter-participant SRM variability decreased, with higher median values and reduced variability at 15° and 20°. Notably, the Individual condition with the same ITD as C3 at 20° (C3 ITD, see Fig. 4) exhibited higher SRM than C3 at that azimuth. The largest differences in median SRM occurred between 5° and 10° across all spatial conditions, with SRM gain decreasing from approximately 10%–15% at 5–10° to 2%–5% at 15–20°.

A two-way repeated-measures ANOVA was considered for both dependent variables (SI and SRM), but normality violations identified via Shapiro–Wilk tests necessitated non-parametric alternatives. Normality was violated for *Spatial Condition = Individual* ($W = 0.946$, $p = 0.016$) and *Azimuth = 20* ($W = 0.903$, $p = 0.007$). For SRM, the only violation was observed for *Azimuth = 20* ($W = 0.876$, $p = 0.001$). Therefore, a Friedman test with the dependent variable of SRM, excluding the co-located condition, was conducted. Effect sizes for Friedman tests are reported as Kendall's W. The test found no significant main effect of *Spatial Condition* ($\chi^2(2) = 0.18$, $p = 0.913$, $W = 0.008$), but a significant

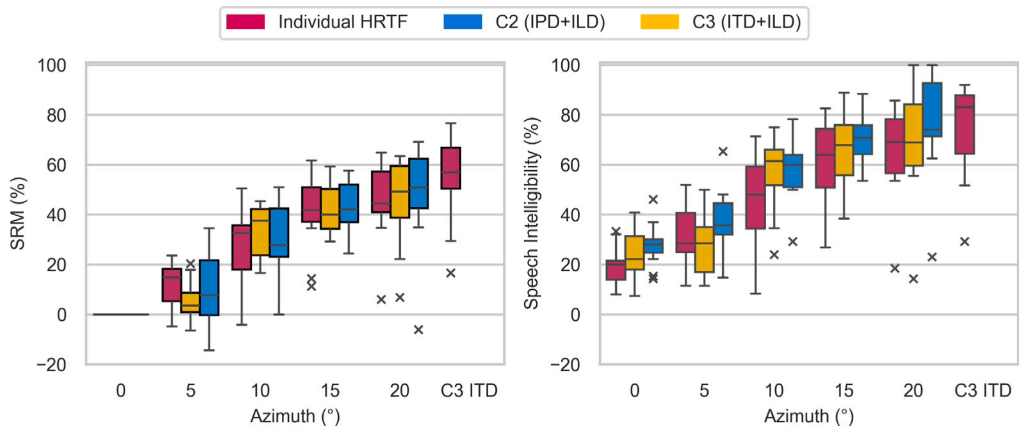

Fig. 4. The left panel shows boxplots of behavioural SI across masker azimuths for individual and magnified HRTFs (C2 & C3). The right panel shows behavioural SRM, calculated as the SI difference between spatially separated and co-located conditions.

main effect of *Azimuth* was observed ($\chi^2(3) = 29.73$, $p < 0.001$, $W = 0.901$). Follow-up Wilcoxon signed-rank tests showed significant differences between all azimuths ($p < 0.01$), with the exception of 15° vs. 20° ($p = 0.054$).

A separate Friedman tests on SI (including the co-located condition) revealed significant main effects of both *Spatial Condition* ($\chi^2(2) = 13.27$, $p = 0.001$, $W = 0.603$) and *Azimuth* ($\chi^2(4) = 40.51$, $p < 0.001$, $W = 0.921$). Post hoc Wilcoxon signed-rank tests with Holm-Bonferroni correction were conducted with effect sizes reported as Rosenthal's r. Post hoc tests indicated significant pairwise differences between *Individual* and *C2* ($W = 0$, $p = 0.003$, $r = 0.912$), and between *C2* and *C3* ($W = 7$, $p = 0.037$, $r = 0.914$), but not between *Individual* and *C3* ($W = 14$, $p = 0.10$, $r = 0.865$). For azimuth, all pairwise comparisons between 0° and 20° were significant ($W = 0$–5, $p < 0.01$, $r = 0.634$–0.966), except 15° versus 20° ($p = 0.054$).

3.3. Behavioural results vs. numerical metrics

Fig. 5A–C illustrate the correlation between behavioural SRM (%) and the predictions derived from numerical measures: LCJ SRM (Fig. 5A), BAM-Q (Fig. 5B), and CSD (Fig. 5C). The C3 ITD data were excluded from analysis to ensure balanced data for each spatial condition. Across all models, correlations were observed between observed and predicted SRM. The BAM-Q model, which predicts binaural quality degradation, yielded the highest correlations among the three models (Individual: r = 0.78, $p < 0.01$; C2: r = 0.82, $p < 0.01$; C3: r = 0.84, $p < 0.01$). The LCJ SRM model also correlated significantly with behavioural SRM across all conditions (Individual: r = 0.61, $p < 0.01$; C2: r = 0.64, $p < 0.01$; C3: r = 0.67, $p < 0.01$). The CSD model exhibited the lowest correlation with SRM for the individual HRTF (Individual: r = 0.49, $p < 0.01$) but higher than LCJ SRM for magnified conditions (C2: r = 0.63, $p < 0.01$; C3: r = 0.77, $p < 0.01$).

Fig. 5D–F present the relationship between predicted and observed SRM gain, which quantifies the benefit of IM relative to individual HRTFs. Unlike the absolute SRM results, none of the metrics showed a significant correlation with behavioural SRM gain. The CSD gain showed a non-significant positive correlation with SRM gain (C3: r = 0.15,p = 0.341; C2: r = 0.04,p = 0.783). Similar non-significance was observed for the BAM-Q gain (C2: r = 0.11,p = 0.409; C3: r = 0.14,p = 0.298), and the LCJ SRM gain (C2: r = −0.07,p = 0.605; C3: r = 0.10,p = 0.456).

At azimuth 0°, although magnification was still applied, the target and masker were processed with the same HRTF. This minimises traditional equalisation-cancellation mechanisms (Durlach, 1963). However, the magnification of interaural cues does change IC, as is further discussed in Section 4. To examine whether changes in IC might relate to the increase in SI observed under co-located conditions for C2 and C3, we analysed the standard deviation of IVS across frequency bands. For all combined target–masker stimuli, in sequential 400-ms windows, we computed the mean IVS as derived in Flessner et al. (2017) within each band for both the reference (Individual HRTF) and test (C2 or C3) conditions. Then we determined the standard deviation of IVS across auditory bands. The mean IVS standard deviation across all stimuli was then calculated per participant and spatial condition. Finally, the difference in IVS standard deviations between the magnified and individual conditions was calculated. This IVS standard deviation difference was plotted against the behavioural differences in SI for all conditions at 0° azimuth, shown in Fig. 6. IVS, as derived in Flessner et al. (2017), was used instead of conventional IC as it accounts for frequency-dependent and reference-level sensitivity of the auditory system.

Fig. 6 shows a significant positive correlation for C2 (r = 0.7, $p < 0.05$), whereas the correlation for C3 did not reach significance (r = −0.16,p = 0.639). This suggests that the increase in SI under co-located conditions, as seen in Fig. 4, may be mediated by magnified IC variance across auditory bands.

3.4. Participant proficiency and SRM gain

Fig. 7 presents the mean SRM gain (%) against behavioural SRM with individual HRTFs stratified into quartiles, for the C2 and C3 conditions, across all masker azimuths. Participants in Q1 (lowest observed SRM with individual HRTFs) exhibited the largest mean SRM gain, with C2 showing a slightly greater effect than C3. A decreasing trend in SRM gain was observed as the observed SRM quartile increased, with participants in Q4 (highest observed SRM with individual HRTFs) showing negative SRM gain. SRM gain was close to zero for individuals in Q2 and Q3 under both conditions.

4. Discussion

4.1. Numerical evaluation

The SRM benefit of each spatial condition was evaluated using the model of Jelfs et al. (2011), which operates on the listener's HRTF and assumes identical target and masker signals. The decrease in SNR in the symmetrical-masker configuration for the magnified conditions can be attributed to the ILD magnification; magnifying ILDs increases ipsilateral HRIR amplitude, thereby reducing the SNR at each ear relative to the target. In practical applications, this limitation may be alleviated by incorporating level compensation into the augmented HRTFs.

A more nuanced effect was observed in the BU component. In condition C2, predicted BU decreased for masker azimuths beyond 25° under both symmetrical and single-masker configurations. This

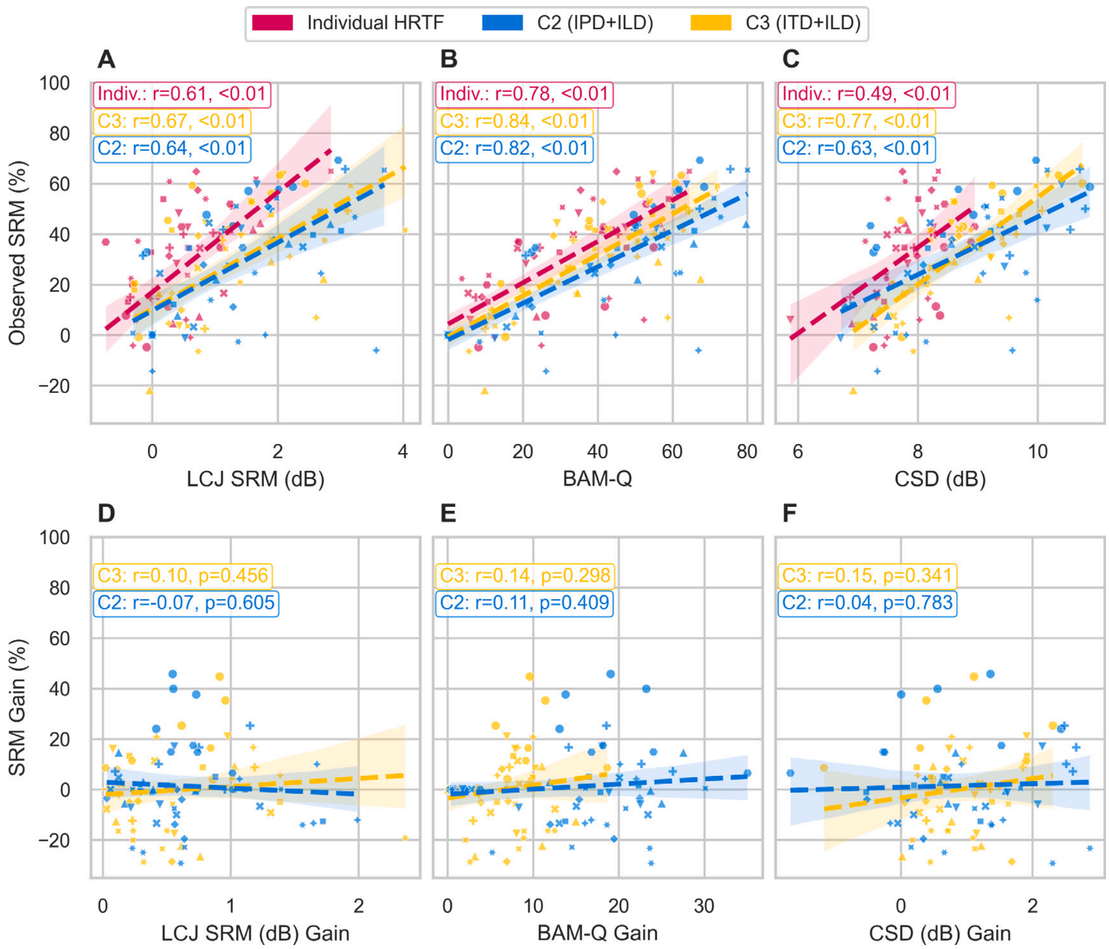

Fig. 5. Panels A–C: Behavioural SRM (%) for individual (Indiv.) and magnified HRTFs (C2&C3) across all masker locations, plotted against numerical predictions (LCJ SRM, BAM-Q, and CSD). Panels D–F: Behavioural SRM differences for C2 and C3 relative to individual HRTF performance, plotted against corresponding differences in numerical metrics. Markers indicate participants. Regression lines with 95% confidence intervals are shown for each spatial condition, annotated with Pearson's r and associated p-values.

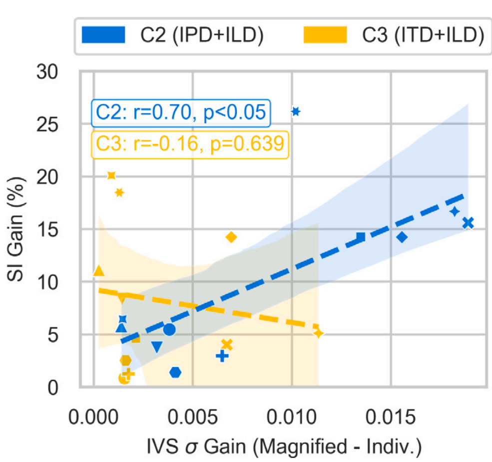

Fig. 6. Correlation between SI under magnified HRTFs (C2 & C3) relative to the individual HRTF (SI gain %) and the difference in standard deviations (σ) of IVS across auditory bands between magnified and individual stimuli.

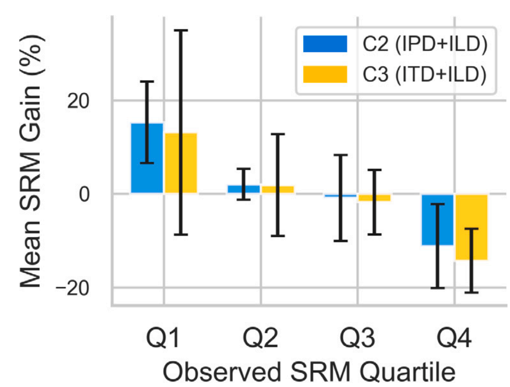

Fig. 7. Behavioural SRM gain (%) stratified by participant performance quartiles with their individual HRTF. Error bars indicate 95% CI.

reduction appears to result from the nonlinear magnification of IPD, which reduces IC, which, according to the equalisation-cancellation theory, negatively impacts BU mechanisms (Durlach, 1963). Thus, while IPD-based magnification increases predicted SRM at smaller azimuths, theoretically, its effectiveness diminishes beyond 25° as IC loss outweighs the unmasking benefit. This pattern was not observed for condition C3, where the application of a constant ITD and hence a linear-phase transformation relative to the individual HRTF did not introduce the same coherence disruptions as C2.

The predicted SRM values for the individual, unmagnified HRTFs from the SONICOM database in our study follow the same trends as Ahrens et al. (2021) using the LISTEN and CIPIC databases, although they differ slightly in predicted values. In the single-masker configuration at 45° azimuth, the median SRM for the CIPIC database is approximately 7.5 dB, and the median for the LISTEN database is slightly higher (≈8 dB). Our predicted SRM is approximately 9 dB, however falls within the range of SRM reported by Ahrens et al. (2021). Our HRTFs exhibited a higher BU (≈3.5 dB at 45° azimuth) than those reported for LISTEN (≈3.2 dB) and CIPIC (≈2 dB), while our SNR component was similar (≈5 dB). These discrepancies are likely due to subject morphology differences and our relatively small sample size. For the symmetrical configuration, our BU was approximately 0.7 dB lower than the median values reported for CIPIC and LISTEN, and our SNR component decreased by approximately 0.5 dB from 0° to ±45°, whereas both CIPIC and LISTEN show an increase of approximately 0.5 dB over a similar range.

Furthermore, a spreading of binaural cues under C2 was observed, consistent with the findings of De Taillez et al. (2018), which is attributed to the nonlinear IPD magnification. In contrast, this spreading is reduced in condition C3 since the ITD magnification ensures a linear-phase transformation relative to the individual HRTF. De Taillez et al. (2018) evaluated cue spreading for an azimuthal target-masker separation of 10° across varying IM exponents and HA directionality algorithms. In the current study, a constant exponent (0.5) was used and spreading was examined as a function of azimuth. The superior dB SRM per degree of spreading observed for C3 relative to both C2 and the individual HRTF supports H2 and suggests that ITD-based IM (C3) is the more favourable approach. However, the differences in SRM are likely below perceptual thresholds, which is further supported by the non-significant behavioural SRM differences between IM conditions.

4.2. Behavioural SRM and SI

Our behavioural results indicate that the IM methods explored did not produce a significant overall enhancement in SRM across the tested azimuths. This outcome is largely attributable to an unexpected increase in SI at 0° for the magnified conditions compared with the Individual HRTF. Although condition C2 yielded significantly higher SI across azimuth, the concomitant increase in SI under co-located conditions meant that the net SRM (computed as the difference between spatially separated and co-located SI) did not differ significantly between conditions. Moreover, the diminishing SI benefit with increasing azimuth, as evidenced by non-significant differences between spatial conditions at 20°, suggests that IM may be most effective when the target and maskers are in close proximity.

Our findings diverge from those of De Taillez et al. (2018), who observed no significant effect of IM on SI using a HA processing approach. Differences in paradigm are likely responsible: their auditory scene consisted of a single masker and target, symmetrically placed around midline with added ‘cafeteria noise’, thereby promoting better–ear listening. In contrast, our setup employed two symmetrical maskers with a central target and no background noise. It is also important to note that their window-based processing may introduce distortions not present in our approach. Furthermore, they hypothesised that the distortion of spatial cues (spreading) may outweigh any SI benefits; however, under co-located conditions, the nonlinear IPD magnification in C2 appears to induce inter-auditory-band correlation variations that may facilitate segregation. Although there was a trend for C3 to increase SI, this effect did not reach significance; a stronger IM factor may be necessary to elicit a significant SI benefit under these conditions.

4.3. Behavioural results vs. numerical metrics

Our analysis of numerical metrics suggests that familiarity with individual cues may influence perceived SRM, consistent with findings by González-Toledo et al. (2024) and Orduña-Bustamante et al. (2018). Listeners tended to perform better with their individual HRTF than with either magnification condition for a given magnitude of interaural cues. Further evidence for the role of familiarity comes from the weak, non-significant correlations between SRM gain and numerical metrics. Neither BU and SNR computations from LCJ SRM nor interaural cue weighting in BAM-Q accounted for SRM variations.

Our results mirror those of Orduña-Bustamante et al. (2018) in which CSD did not account for SRM differences between individual and KEMAR HRTFs. Our CSD differences were also not significantly correlated with SRM gain. However, CSD was correlated with the behavioural SRM across masker separations, suggesting CSD may be a useful metric for predicting absolute SRM. Nevertheless, comparing behavioural SRM % against dB predictions from LCJ SRM may have caused low correlations, especially since our correlation was lower than Cuevas-Rodriguez et al. (2021), who compared observed speech reception thresholds in dB, with model predictions. Taken together, this supports H3, suggesting that numerical signal differences are not the sole factor determining SRM gain.

When the target and maskers were co-located, a notable phenomenon was observed: SI was higher under C3, and especially C2, compared to the individual HRTF condition. It has been shown that the auditory system is sensitive to IC variations in narrow bands (Koehnke et al., 1986; Jain et al., 1991). Speech perception depends on the relative strengths of harmonics that the auditory system can distinguish, particularly in the frequency regions where vowel formants occur. Consequently, Culling et al. (2001) suggests that if listeners can discriminate different degrees of interaural decorrelation, they might be able to infer the relative intensity of components of speech (e.g. formants) by analysing how much they reduce the interaural correlation in specific frequency bands. Consequently, they suggest that listeners might use the spectral distribution of IC as a perceptual cue to identify speech elements that would otherwise be masked. Listeners were also especially sensitive to changes at higher reference correlation levels.

Under co-located conditions, IC is inherently high due to minimal interaural disparities in the HRTF, making listeners more perceptually attuned to subtle variations in IC. This frequency- and reference-dependent IC sensitivity is integrated into the IVS calculation of the BAM-Q model.

In light of this, we tentatively hypothesise that listeners may be exploiting IC variations across auditory bands to perceptually segregate targets from maskers, thereby explaining the observed SI trend (C2 (IPD+ILD) > C3 (ITD+ILD) > Indiv). Instead of applying a constant time-domain delay, C2 applies IPD exponentiation. This results in larger IC variations across auditory bands, particularly noticeable in the 1–2 kHz range, where IVS sensitivity is highest (Flessner et al., 2017). Given that target and masker content spans different spectral regions, listeners may exploit these IC variations to unmask the target.

Conversely, C3 applies a constant time-domain delay to the contralateral ear – a linear-phase operation that preserves the relative phase structure of each ear with respect to the original HRTF. As a result, IC variations across frequency bands are not exaggerated relative to the individual HRTF, providing fewer IC-based segregation cues. Nevertheless, in both C2 and C3, ILD magnification introduces slight asymmetries even at 0° azimuth, which may account for the trend that C3 outperforms the unmodified individual HRTF.

The significant correlation between IVS standard deviation differences and SI between C2 and individual HRTF stimuli seems to support this hypothesis. As such, the spatial cue spreading found here, also reported by De Taillez et al. (2018), may be beneficial under co-located target-masker conditions.

4.4. Participant proficiency and SRM gain

The results depicted in Fig. 7 reveal a non-linearity in the effect of IM on SRM gain when stratified by baseline performance. Our findings suggest that for listeners already achieving robust SRM with their individual spatial cues, further magnification may be detrimental. These results are analogous to findings in other studies, although the experimental paradigms differ. For instance, a study on spatial auditory training for speech perception in noise found that participants with poorer initial performance tended to show greater improvement (De Boer and Thornton, 2008). Also, in the study of Kollmeier and Peissig (1990), only HI who already had some baseline binaural intelligibility advantage showed improvements with IM.

4.5. Future work

The findings indicate a significant effect of magnification on SI for the IPD and ILD (C2) condition and a trend towards increased SI under C3; however, no significant increase in SRM. Furthermore, as was found by De Taillez et al. (2018), C2 causes spatial spreading, which we show increases with azimuth from auditory model predictions. Furthermore, the indication of an offset in SI under individual listening compared with C2 and C3 indicates an aspect of listener familiarity with the spatial cues.

This relative SI offset, along with spatial distortions of magnification, raises the question of whether we can train with an augmented HRTF to make better use of the enhanced spatial cues and minimise excessive lateralisation and spreading. The human auditory system is demonstrably plastic: listeners can remap an HRIR to a novel location (Shinn-Cunningham et al., 1998) and re-weight inconsistent ITD/ILD cues (Klingel et al., 2021). Furthermore, after spectral-cue manipulations via ear-moulds, localisation accuracy recovers to baseline within three-six weeks (Hofman et al., 1998), with daily sensory-motor training accelerating this adaptation to near baseline in approximately ten days (Carlile et al., 2014). Task-specific spatial-hearing training has likewise been shown to enhance SI (Amiri et al., 2022). Taken together, these studies suggest that a paradigm may be devised to assess the effect of IM on SI and SRM after training. For instance, a multi-session programme with audio-visual feedback that maps the lateralised locations generated by C2 or C3 back to their veridical azimuths, thereby retaining the enlarged interaural cues, may suppress lateral distortion. Alternatively, an adaptive speech-reception-threshold task similar to Amiri et al. (2022), in which participants practise target recognition amid magnified-cue maskers, progressively familiarising themselves with the altered cue structure, may be used. In this case, participants who trained with the magnified HRTFs could be hypothesised to improve more than those training with unmagnified individual HRTFs.

Several limitations of this work should be addressed in future studies. The relatively small sample size constrains the generalisability of the results. Furthermore, the augmented conditions introduced unfamiliar spatial cues, which, alongside a lack of acoustic simulation, may have caused internalised auditory percepts. Consequently, future work should explore more complex auditory environments to allow for ecologically valid assessments of IM.

5. Conclusion

In this study, we investigated the feasibility of enhancing SRM through direct augmentation of individual HRTFs. We evaluated two HRTF augmentation approaches: combining IPD and ILD magnification (C2) and another applying a constant time delay alongside ILD magnification (C3). Behaviourally, magnification did not significantly impact SRM. However, C2 led to significantly higher SI compared to listeners' own HRTFs, whereas C3 did not. Auditory model results indicated that C2 decreased IC with masker azimuth, mitigating BU effects at more lateral locations, with C3 providing a better trade-off between spatial cue spreading and SRM, particularly at larger masker azimuths. Under co-located target and maskers, IPD magnification of C2 amplifies inter-auditory-band IC variation, which we suggest facilitates SI. This SI increase in co-located target-masker configurations may explain the lack of a significant effect of magnification on SRM. Additionally, our findings suggest that individuals with lower behavioural SRM using their own HRTFs benefit more from augmentation. Behavioural comparisons with auditory model predictions also suggest that listener familiarity with individual spatial cues may influence the effectiveness of IM. Future work should refine HRTF augmentation algorithms to minimise perceptual distortions, and investigate whether training with augmented HRTFs improves the utilisation of magnified cues.

CRediT authorship contribution statement

Nils Marggraf-Turley: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Niels H. Pontoppidan:** Supervision. **Lorenzo Picinali:** Supervision.

Ethics approval

This study was approved by the Research Governance and Integrity Team (RGIT) at Imperial College London. SETREC No. 7046527. Informed consent was obtained from all participants, and the privacy rights of subjects were observed.

Funding sources

This work is funded by an EPSRC Case Conversion scholarship and co-funded by Eriksholm Research Centre, Oticon A/S.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix. Magnified stimuli generation

Steps marked with [C2&3] are common to both magnification methods, while those marked with [C2] or [C3] indicate operations specific to Condition 2 or Condition 3, respectively.

[C2&3]: For a given target or masker location, the corresponding individual HRIRs were selected. These were transformed into the frequency domain via discrete Fourier transforms, yielding complex spectra for each ear, which were decomposed into magnitude and phase components:

$$R[k] = |R[k]|\, e^{i\phi_R[k]}, \quad L[k] = |L[k]|\, e^{i\phi_L[k]}, \tag{A.1}$$

where $|R[k]|$ and $|L[k]|$ are the magnitude spectra, and $\phi_R[k]$ and $\phi_L[k]$ are the corresponding phase spectra for the right and left ears, respectively. For each frequency bin k, the modified spectra are defined as follows, with k_c denoting the crossover bin corresponding to 1.5 kHz, and α is the magnification exponent, set to 0.5:

C2 IPD Magnification (for $k \le k_c$):

$$\hat{R}[k] = |R[k]|\, e^{i\phi_R[k]} \cdot e^{i\alpha[\phi_R[k]-\phi_L[k]]}, \tag{A.2}$$

$$\hat{L}[k] = |L[k]|\, e^{i\phi_L[k]} \cdot e^{i\alpha[\phi_L[k]-\phi_R[k]]}, \tag{A.3}$$

This step magnifies the low-frequency localisation cues derived from timing differences.

C2&3 ILD Magnification (for $k > k_c$):

$$\hat{R}[k] = |R[k]| \left[\frac{|R[k]|+\epsilon}{|L[k]|+\epsilon}\right]^{\alpha} e^{i\phi_R[k]}, \tag{A.4}$$

$$\hat{L}[k] = |L[k]| \left[\frac{|L[k]|+\epsilon}{|R[k]|+\epsilon}\right]^{\alpha} e^{i\phi_L[k]}, \tag{A.5}$$

where ϵ is a small constant (e.g., 10^{-6}) added to prevent division by zero. This step magnifies intensity-based directional cues by applying a nonlinear amplification of amplitude asymmetry between ears.

[C2&3]: The modified HRIRs were then obtained by applying the inverse discrete Fourier transform:

$$\hat{r}_R[n] = \mathcal{F}^{-1}\{\hat{R}[k]\}, \quad \hat{r}_L[n] = \mathcal{F}^{-1}\{\hat{L}[k]\}. \tag{A.6}$$

[C3]: Instead of magnifying phase directly, C3 applies a time-domain delay to the contralateral ear. First, the ITD was estimated between the original HRIRs by finding the lag that gives the highest cross-correlation between the HRIRs.

$$\mathrm{ITD} = \frac{1}{f_s} \cdot \arg\max_{\tau} \left[(r_R \star r_L)[\tau] \right], \tag{A.7}$$

where $\star$ denotes discrete time cross-correlation. The contralateral impulse response is then shifted by a magnified version of the original ITD:

$$\hat{r}_{\mathrm{contra}}[n] = \hat{r}_{\mathrm{contra}}[n + \Delta n], \quad \text{where } \Delta n = \mathrm{round}(\alpha \cdot \mathrm{ITD} \cdot f_s), \tag{A.8}$$

[C2&3]: Binaural stimuli were generated by convolving the magnified HRIRs with the stimulus signal $x[n]$:

$$s_R[n] = \hat{r}_R[n] * x[n], \quad s_L[n] = \hat{r}_L[n] * x[n], \tag{A.9}$$

where $*$ denotes discrete-time convolution.

Data availability

The behavioural data that support the findings of this study are available on the following Zenodo repository: 10.5281/zenodo.15198918.

References

Ahrens, A., Cuevas-Rodriguez, M., Brimijoin, W.O., 2021. Speech intelligibility with various head-related transfer functions: A computational modelling approach. JASA Express Lett. 1, 034401. http://dx.doi.org/10.1121/10.0003618.

Amiri, M., Jarollahi, F., Jalaie, S., Sameni, S.J., 2022. Auditory spatial training-induced changes in the release of informational masking in older adults. Audit. Vestib. Res. 31, http://dx.doi.org/10.18502/avr.v31i4.10735.

Armstrong, C., Thresh, L., Murphy, D., Kearney, G., 2018. A perceptual evaluation of individual and non-individual HRTFs: A case study of the SADIE II database. Appl. Sci. 8, 2029. http://dx.doi.org/10.3390/app8112029.

Bolia, R.S., Nelson, W.T., Ericson, M.A., Simpson, B.D., 2000. A speech corpus for multitalker communications research. J. Acoust. Soc. Am. 107, 1065–1066. http://dx.doi.org/10.1121/1.428288.

Brown, C.A., 2014. Binaural enhancement for bilateral cochlear implant users. Ear Hear. 35, 580–584. http://dx.doi.org/10.1097/AUD.0000000000000044.

Brungart, D.S., Romigh, G.D., 2009. Spectral HRTF enhancement for improved vertical-polar auditory localization. In: 2009 IEEE Workshop on Applications of Signal Processing To Audio and Acoustics. IEEE, New Paltz, NY, USA, pp. 305–308. http://dx.doi.org/10.1109/ASPAA.2009.5346479.

Brungart, D.S., Simpson, B.D., 2009. Effects of bandwidth on auditory localization with a noise masker. J. Acoust. Soc. Am. 126, 3199–3208. http://dx.doi.org/10.1121/1.3243309.

Carlile, S., Balachandar, K., Kelly, H., 2014. Accommodating to new ears: The effects of sensory and sensory-motor feedback. J. Acoust. Soc. Am. 135, 2002–2011. http://dx.doi.org/10.1121/1.4868369.

Cherry, E.C., 1953. Some experiments on the recognition of speech, with one and with two ears. J. Acoust. Soc. Am. 25, 975–979. http://dx.doi.org/10.1121/1.1907229.

Cuevas-Rodriguez, M., Gonzalez-Toledo, D., Reyes-Lecuona, A., Picinali, L., 2021. Impact of non-individualised head related transfer functions on speech-in-noise performances within a synthesised virtual environment. J. Acoust. Soc. Am. 149, 2573–2586. http://dx.doi.org/10.1121/10.0004220.

Culling, J.F., Colburn, H.S., Spurchise, M., 2001. Interaural correlation sensitivity. J. Acoust. Soc. Am. 110, 1020–1029. http://dx.doi.org/10.1121/1.1383296.

Daugintis, R., Alary, B., Geronazzo, M., Picinali, L., 2024. Effects of binaural rendering personalisation and reverberation on speech-on-speech masking. In: AES 5th International Conference on Audio for Virtual and Augmented Reality, AES. Redmond, WA, United States.

De Boer, J., Thornton, A.R.D., 2008. Neural correlates of perceptual learning in the auditory brainstem: efferent activity predicts and reflects improvement at a speech-in-noise discrimination task. J. Neurosci. 28, 4929–4937. http://dx.doi.org/10.1523/JNEUROSCI.0902-08.2008.

De Taillez, T., Grimm, G., Kollmeier, B., Neher, T., 2018. Acoustic and perceptual effects of magnifying interaural difference cues in a simulated binaural hearing aid. Int. J. Audiol. 57, S81–S91. http://dx.doi.org/10.1080/14992027.2017.1308564.

Dietz, M., Ewert, S.D., Hohmann, V., 2011. Auditory model based direction estimation of concurrent speakers from binaural signals. Speech Commun. 53, 592–605. http://dx.doi.org/10.1016/j.specom.2010.05.006.

Durlach, N.I., 1963. Equalization and cancellation theory of binaural masking-level differences. J. Acoust. Soc. Am. 35, 1206–1218. http://dx.doi.org/10.1121/1.1918675.

Durlach, N.I., Pang, X.D., 1986. Interaural magnification. J. Acoust. Soc. Am. 80, 1849–1850. http://dx.doi.org/10.1121/1.394252.

Engel, I., Daugintis, R., Vicente, T., Hogg, A.O.T., Pauwels, J., Tournier, A.J., Picinali, L., 2023. The SONICOM HRTF dataset. J. Audio Eng. Soc. 71, 241–253. http://dx.doi.org/10.17743/jaes.2022.0066.

Flessner, J.H., Huber, R., Ewert, S., 2017. Assessment and prediction of binaural aspects of audio quality. J. Audio Eng. Soc. 65, 929–942. http://dx.doi.org/10.17743/jaes.2017.0037.

González-Toledo, D., Cuevas-Rodríguez, M., Vicente, T., Picinali, L., Molina-Tanco, L., Reyes-Lecuona, A., 2024. Spatial release from masking in the median plane with nonnative speakers using individual and mannequin head related transfer functions. J. Acoust. Soc. Am. 155, 284–293. http://dx.doi.org/10.1121/10.0024239.

Guiraud, P., Sum, K., Pontoppidan, N., Poole, K., Picinali, L., 2024. Adaptation to altered interaural time differences in a virtual reality environment. In: Proceedings of the 10th Convention of the European Acoustics Association Forum Acusticum 2023. European Acoustics Association, Turin, Italy, pp. 1773–1780. http://dx.doi.org/10.61782/fa.2023.0412.

Hofman, P.M., Van Riswick, J.G.A., Van Opstal, A.J., 1998. Relearning sound localization with new ears. Nature Neurosci. 1, 417–421. http://dx.doi.org/10.1038/1633.

Jain, M., Gallagher, D.T., Koehnke, J., Colburn, H.S., 1991. Fringed correlation discrimination and binaural detection. J. Acoust. Soc. Am. 90, 1918–1926. http://dx.doi.org/10.1121/1.401671.

Jelfs, S., Culling, J.F., Lavandier, M., 2011. Revision and validation of a binaural model for speech intelligibility in noise. Hear. Res. 275, 96–104. http://dx.doi.org/10.1016/j.heares.2010.12.005.

Kaneko, S., Suenaga, T., Sekine, S., 2016. DeepEarNet: Individualizing spatial audio with photography, ear shape modeling, and neural networks. J. Audio Eng. Soc..

Klingel, M., Kopčo, N., Laback, B., 2021. Reweighting of binaural localization cues induced by lateralization training. JARO 22, 551–566. http://dx.doi.org/10.1007/s10162-021-00800-8.

Koehnke, J., Colburn, H.S., Durlach, N.I., 1986. Performance in several binaural-interaction experiments. J. Acoust. Soc. Am. 79, 1558–1562. http://dx.doi.org/10.1121/1.393682.

Kollmeier, B., Peissig, J., 1990. Speech intelligibility enhancement by interaural magnification. Acta Oto-Laryngol. 109, 215–223. http://dx.doi.org/10.1080/00016489.1990.12088432.

Lavandier, M., Culling, J.F., 2010. Prediction of binaural speech intelligibility against noise in rooms. J. Acoust. Soc. Am. 127, 387–399. http://dx.doi.org/10.1121/1.3268612.

Lopez-Poveda, E.A., Eustaquio-Martín, A., Stohl, J.S., Wolford, R.D., Schatzer, R., Wilson, B.S., 2016. A binaural cochlear implant sound coding strategy inspired by the contralateral medial olivocochlear reflex. Ear Hear. 37, e138–e148. http://dx.doi.org/10.1097/AUD.0000000000000273.

Majdak, Piotr, Hollomey, Clara, Baumgartner, Robert, 2022. AMT 1.x: A toolbox for reproducible research in auditory modeling. Acta Acust. 6 (19), http://dx.doi.org/10.1051/aacus/2022011.

Miccini, R., Spagnol, S., 2020. HRTF individualization using deep learning. In: 2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops. VRW, pp. 390–395. http://dx.doi.org/10.1109/VRW50115.2020.00084.

Middlebrooks, J.C., 1999. Virtual localization improved by scaling nonindividualized external-ear transfer functions in frequency. J. Acoust. Soc. Am. 106, 1493–1510. http://dx.doi.org/10.1121/1.427147.

Morikawa, D., Hirahara, T., 2010. Signal frequency range necessary for horizontal sound localization. Acoust. Sci. Technol. 31, 417–419. http://dx.doi.org/10.1250/ast.31.417.

Orduña-Bustamante, F., Padilla-Ortiz, A., Torres-Gallegos, E.A., 2018. Binaural speech intelligibility through personal and non-personal HRTF via headphones, with added artificial noise and reverberation. Speech Commun. 105, 53–61. http://dx.doi.org/10.1016/j.specom.2018.10.009.

Pirhosseinloo, S., Kokkinakis, K., 2016. An interaural magnification algorithm for enhancement of naturally-occurring level differences. In: Interspeech 2016. ISCA, Vol. 255, pp. 8–2561. http://dx.doi.org/10.21437/Interspeech.2016-1049.

Richardson, B.N., Kainerstorfer, J.M., Shinn-Cunningham, B.G., Brown, C.A., 2025. Magnified interaural level differences enhance binaural unmasking in bilateral cochlear implant users. J. Acoust. Soc. Am. 157 2, 1045–1056. http://dx.doi.org/10.1121/10.0034869.

Shinn-Cunningham, B.G., Durlach, N.I., Held, R.M., 1998. Adapting to upernormal auditory localization cues. I. bias and resolution. J. Acoust. Soc. Am. 103, 3656–3666. http://dx.doi.org/10.1121/1.423088.

So, R.H.Y., Leung, N.M., Horner, A.B., Braasch, J., Leung, K.L., 2011. Effects of spectral manipulation on nonindividualized head-related transfer functions (HRTFs). Hum. Factors: J. Hum. Factors Ergon. Soc. 53, 271–283. http://dx.doi.org/10.1177/0018720811406883.

Yao, S.N., Chen, L.J., 2013. HRTF adjustments with audio quality assessments. Arch. Acoust. 38, 55–62. http://dx.doi.org/10.2478/aoa-2013-0007.