


The Role of Interaural Differences, Head Shadow, and Binaural Redundancy in Binaural Intelligibility Benefits Among School-Aged Children

Trends in Hearing
Volume 25: 1–14
© The Author(s) 2021
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/23312165211045313
journals.sagepub.com/home/tia


Z. Ellen Peng¹  and Ruth Y. Litovsky¹ 

Abstract

In complex listening environments, children can benefit from auditory spatial cues to understand speech in noise. When a spatial separation is introduced between the target and masker and/or listening with two ears versus one ear, children can gain intelligibility benefits with access to one or more auditory cues for unmasking: monaural head shadow, binaural redundancy, and interaural differences. This study systematically quantified the contribution of individual auditory cues in providing binaural speech intelligibility benefits for children with normal hearing between 6 and 15 years old. In virtual auditory space, target speech was presented from +90° azimuth (i.e., listener's right), and two-talker babble maskers were either co-located (+90° azimuth) or separated by 180° (−90° azimuth, listener's left). Testing was conducted over headphones in monaural (i.e., right ear) or binaural (i.e., both ears) conditions. Results showed continuous improvement of speech reception threshold (SRT) between 6 and 15 years old and immature performance at 15 years of age for both SRTs and intelligibility benefits from more than one auditory cue. With early maturation of head shadow, the prolonged maturation of unmasking was likely driven by children's poorer ability to gain full benefits from interaural difference cues. In addition, children demonstrated a trade-off between the benefits from head shadow versus interaural differences, suggesting an important aspect of individual differences in accessing auditory cues for binaural intelligibility benefits during development.

Keywords

spatial release from masking, binaural squelch, head shadow, interaural differences, auditory development

Received 21 August 2020; Revised 19 August 2021; accepted 20 August 2021

Introduction

For children, the ability to segregate sounds from background noise is crucial for successful communication in complex auditory environments such as noisy classrooms. When noise sources are located away from the target speech, listeners have access to multiple auditory cues from the spatial separation to improve intelligibility of the target speech (Blauert, 1997; Bronkhorst, 2015; Plomp, 1976). In this study, we assessed the individual contribution of auditory cues in providing intelligibility benefits among children with normal hearing (NH), with a focus on understanding the developmental trajectory.

The intelligibility benefit from spatial separation is known as a spatial release from masking (SRM), which is typically computed as the difference in percent correct scores or in speech reception thresholds (SRTs) between conditions in

which the target and masker are co-located versus spatially separated. SRM mainly arises from two auditory cues: monaural head shadow and interaural differences, namely interaural time and level differences (ITD and ILD) (Bronkhorst, 2015). Studies with children typically positioned the target at 0° azimuth (i.e., front of listener), with spatial separation either symmetrical or asymmetrical (e.g., Misurelli and Litovsky, 2012). Asymmetrical spatial separation provides access to both interaural differences and monaural head shadow cues. Monaural head shadow here specifically refers

¹Waisman Center, University of Wisconsin-Madison, Madison, WI, USA

Corresponding Author:

Z. Ellen Peng, Boys Town National Research Hospital, 555 N. 30th Street, Omaha, NE 68131.
Email: ellen.peng@boystown.org



to a better signal-to-noise ratio (SNR) of the frontal target in the ear contralateral (further) to the masker as compared to the ear ipsilateral (closer) to the masker. Even though the auditory cue is the result of the spectrally dependent ILD, the head shadow is a monaural listening phenomenon. When maskers are displaced symmetrically, the monaural head shadow is minimized, resulting in interaural differences as the more useful cues for SRM.

There has been a considerable amount of research examining SRM development in children (Yuen & Yuan, 2014 for a review). When both monaural head shadow and interaural differences are available, several studies show that SRM emerges for children as young as 3–4 years old (Ching et al., 2011; Garadat & Litovsky, 2007; Johnstone & Litovsky, 2006; Litovsky, 2005). Across studies, NH children demonstrate a 3–11 dB SRM (Ching et al., 2011; Garadat & Litovsky, 2007; Litovsky, 2005; Misurelli & Litovsky, 2012, 2015; Van Deun et al., 2010; Yuen & Yuan, 2014), as compared to a 5–12 dB SRM in NH adults (Hawley et al., 2004; Ihlefeld & Shinn-Cunningham, 2008; Jones & Litovsky, 2008). Using monosyllabic or disyllabic words as the target speech, several studies showed SRM maturation as early as around 5 years old (Litovsky, 2005; Misurelli & Litovsky, 2012; Murphy et al., 2011). Others using sentences as the target speech showed a later age of maturation than studies using target words. When the masker is more distinct from the target, such as using child babble to mask adult speech, children reached SRM maturity by 6 years of age (Griffin et al., 2019). When the target and masker are more similar, such as that both the target and maskers are of same-sex or same-voice, SRM maturation is prolonged to 9 years or older (Brown et al., 2010; Cameron et al., 2011; Corbin et al., 2017; Vaillancourt et al., 2008).

Multiple studies have compared access to monaural head shadow and interaural differences for SRM in children. To directly measure the head shadow effect in SRM for children, Van Deun et al. (2010) and Corbin et al. (2017) used a blocked-ear method with earplugs to create monaural listening in free-field. An averaged 3–4 dB SRM from head shadow was reported for NH children between 4 and 10 years old (Corbin et al., 2017; Van Deun et al., 2010). Both studies also included test conditions with binaural listening when the target and maskers were spatially separated. The SRM from interaural differences was derived at ~0 dB for the 4–8-year-old children (Van Deun et al., 2010), but larger at ~4 dB for 8–10-year-old children (Corbin et al., 2017). Others have used a symmetrical spatial separation to directly measure SRM from interaural differences. The Listening in Spatialized Noise-Sentences Test (LiSN-S) developed by Cameron and Dillon (2007) showed an averaged of 9–11 dB SRM from interaural differences in children between 6 and 11 years old. By testing both the asymmetrical and symmetrical separations in the same children in two age groups between 4 and 7 years old on average, Misurelli and Litovsky (2012) showed that children achieved larger SRM

in the asymmetrical separation (i.e., 5–7 dB on average) than in the symmetrical separation (i.e., 3–5 dB on average), and suggested adult-like SRM before 5 years of age. They further showed an averaged 2–3 dB SRM from the monaural head shadow by subtracting symmetrical SRM from asymmetrical SRM. In a follow-up study, Misurelli and Litovsky (2015) compared SRMs measured using same- versus different-sex two-talker masker (cf., target talker) and found that minimizing voice cues improved SRM, but it had no impact on SRM development. Griffin et al. (2019) examined a larger age range of NH children between 6 and 12 years old on both symmetrical and asymmetrical SRMs using a two-talker child babble and sentences as target speech. They reported an average 5 dB asymmetrical SRM and a smaller average 1–3 dB symmetrical SRM, leading to an estimated 2–4 dB benefit from monaural head shadow.

A third auditory cue known as binaural redundancy might also contribute to binaural speech intelligibility benefits. Compared to listening to one ear, listening with both ears not only provides access to interaural differences, but also the access to coherent signals from the second ear (Bronkhorst & Plomp, 1988; Epstein & Florentine, 2009). When speech is presented in quiet, binaural redundancy provides 2–3 dB intelligibility benefit in NH adults (Epstein & Florentine, 2009). For speech in noise, binaural redundancy provides a benefit between 1 and 3 dB, depending on the masker type (Bronkhorst & Plomp, 1988; Dieudonné & Francart, 2019). Binaural redundancy is often reported as part of the binaural squelch effect, in which adding an ear with a poorer target SNR provides an intelligibility benefit as compared to monaural listening (Bronkhorst, 2015). Few studies have reported binaural squelch and binaural redundancy in NH listeners due to the difficulty in effectively simulating a monaural listening condition in free-field testing (Van Deun et al., 2010). By using virtual auditory space (VAS), Dieudonné and Francart (2019) tested NH adults with spatialized sounds delivered over headphones; monaural listening was created more successfully than in free-field by muting audio input to one ear. Using a single-talker masker, they found that binaural redundancy (~2–3 dB) contributed more to binaural intelligibility benefits than interaural differences (< 1 dB) for NH adults.

Much less is known regarding how children access individual auditory cues for binaural intelligibility benefits when all three cues are available, namely, head shadow, interaural differences, and binaural redundancy. Existing literature in children has so far focused on parsing SRM due to interaural differences from monaural head shadow (Griffin et al., 2019; Misurelli & Litovsky, 2012, 2015), with limited evidence to suggest that children receive binaural intelligibility benefits from binaural redundancy (Van Deun et al., 2010). A more comprehensive understanding of NH children's access to individual auditory cues for binaural intelligibility benefits will provide a more detailed basis for

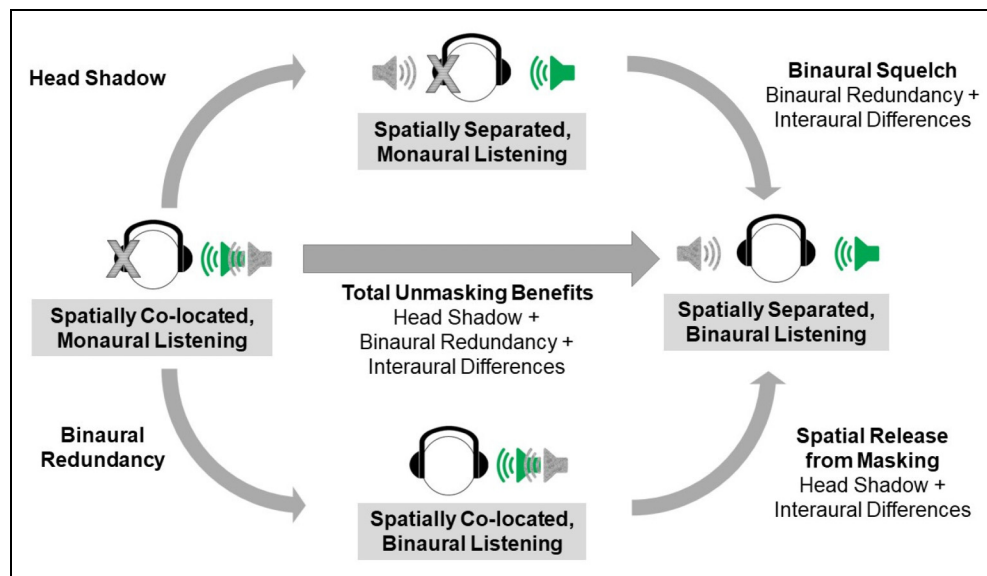


Figure 1. Schematics showing four test conditions to measure intelligibility benefits from auditory cues that contribute to spatial unmasking: head shadow, binaural redundancy, and interaural differences. The effects of spatial release from masking, binaural squelch, and spatial unmasking are the benefits from using multiple cues. Arrows indicate an improvement in speech intelligibility by comparing each pair of test conditions. Loudspeaker symbols indicate virtual locations of target (green) at $+90^\circ$ azimuth and maskers (gray) at $+90^\circ$ azimuth for spatially co-located condition and -90° azimuth for 180° spatial separation. Adapted from Peng & Litovsky, 2021.

fitting bilateral hearing devices, such as hearing aids (HAs) and cochlear implants (CIs) for children with hearing loss, with the potential to maximize access to auditory cues for optimal speech perception in noise.

The present study aimed at characterizing the developmental change in intelligibility benefits from individual auditory cues. We used a fixed angular separation of 180° , which provided maximum access to head shadow and interaural difference cues. NH children ranging in age from 6 to 15 years were studied. The contribution of individual auditory cues or a combination of more than one cue was computed by comparing SRTs across several key conditions. In addition to examining the age effect, we hypothesized that NH children would demonstrate larger intelligibility benefits than previously reported in other studies that used a 90° angular separation due to the enlarged auditory cues and separating the target and masker in opposing hemifields (Ching et al., 2011; Misurelli & Litovsky, 2012; Van Deun et al., 2010).

As illustrated in Figure 1, we used a theoretical framework similar to that in Dieudonné and Francart (2019) to measure SRTs in four test conditions. All stimuli were created in VAS using head-related transfer functions (HRTFs) from a standard manikin and presented over headphones. We avoided the blocked ear method used in Van Deun et al. (2010) and free-field testing because residual audibility from earplugs may bias the measurement of binaural intelligibility benefits from access to individual auditory cues. Further, VAS-based measurement of SRM has been demonstrated in both NH adults and children. The LiSN-S test is the only one that was developed for children between 6 and 11 years old

using VAS with HRTFs from a standard manikin (Cameron & Dillon, 2007). Several studies supporting the LiSN-S test development showed that children could use interaural differences cues from a standard manikin for SRM, even though these cues do not perfectly match their own (Brown et al., 2010; Cameron et al., 2006; Cameron et al., 2011).

Materials and Methods

Participants

Participants included 31 children and 10 adults. Children were between 6.9 and 15.7 years old recruited from three age groups: 6–9 ($n = 10$, $M = 8.6$ years, $SD = 1.1$), 10–12 ($n = 11$, $M = 11.6$ years, $SD = 0.8$), and 13–15 ($n = 10$, $M = 13.9$ years, $SD = 0.9$). All children were typically developing with no known developmental delays, including hearing or speech impairments. On the day of testing, none of the children had a known illness or ear infection based on parental reports. All adult participants were recruited on the campus of the University of Wisconsin-Madison and were between 18.2 and 25.9 years of age ($M = 21.7$ years, $SD = 2.4$). All participants had pure-tone hearing thresholds at or below 20 dB hearing levels in both ears at octave band center frequencies between 125 and 8000 Hz.

All experimental protocols and procedures were approved by the Health Sciences Institutional Review Board at the University of Wisconsin-Madison. Informed consent was obtained from each child's parent or legal guardian, along

with assents from the child prior to testing. Children were compensated \$7.50/h in addition to multiple small prizes during the test session. All adult participants provided written consent and received compensation of \$8.00/h.

Speech Materials

Target speech was open-set sentences from the Australian Speech in Noise Test (AuSTIN; Dawson et al., 2013). We chose open-set sentences to create a task that better reflects everyday verbal communication for children in critical learning environments. Open-set sentence recognition is also more sensitive to developmental effects of speech perception in noise than word recognition (Cameron et al., 2006; Corbin et al., 2016; Griffin et al., 2019). The AuSTIN corpus contains over 1200 Bamford-Kowal-Bench-like sentences suitable for speech intelligibility testing with children as young as 6 years of age. As an example, a target sentence is “the LOCKED the CAR DOOR” with the three keywords capitalized. Some of the sentences were slightly modified to contain familiar vocabulary for American English-speaking children. A subset of 597 AuSTIN sentences containing three keywords was used with keyword-based scoring. The AuSTIN sentences were recorded by a 26-year-old female native talker of American English, who was from the Midwest but with a minimal regional accent. The target talker has an averaged fundamental frequency of 195 Hz.

The masker was two-talker babble consisting of continuous discourse, specifically age-appropriate short science stories (TIME USA, n.d.). Previous work has shown that two-talker speech has a greater masking effect than a single-talker speech by limiting opportunities to “glimpse” through silence gaps (Buss et al., 2017), as well as a greater amount of “informational masking” than maskers with more than two talkers (Kidd et al., 2016). In this study, the masker speech materials were recorded by a second female talker. The masker talker was a 22-year-old female from the Midwest with a minimal regional accent; she has an average fundamental frequency of 251 Hz that was slightly higher than that of the target female talker. Both talkers had a similar speech rate at 3.5 syllables/s. By using a two-talker same-sex masker which maximized informational masking, listeners were expected to mostly rely on auditory spatial cues for unmasking (Brungart et al., 2001; Cameron & Dillon, 2007; Freyman et al., 2001; Johnstone & Litovsky, 2006; Leibold et al., 2018; Leibold et al., 2020; Misurelli & Litovsky, 2015).

Auditory Stimuli

Auditory stimuli with spatial perception were created in VAS using HRTFs measured from a standard KEMAR manikin (GRAS Sound & Vibration, Holte, Denmark). In VAS, the target was always positioned at $+90^\circ$ azimuth or on the side closer to the listener’s right ear. The masker was placed either at the same spatial location as the target (i.e.,

spatially co-located) or at -90° azimuth on the side closer to the listener’s left ear (i.e., spatially separated). When the masker was located at -90° azimuth, a 180° spatial separation was introduced which maximized both the head shadow and interaural difference cues to distinguish the target and masker streams. All speech materials were digitally convolved with the KEMAR HRTFs measured at these two spatial locations and presented over headphones.

To calibrate playback intensity in VAS, we chose the HRTFs recorded at the 0° azimuthal angle, convolved with the target speech, and calibrated the long-term averaged root-mean-squared (RMS) intensity at 60 dB SPL (A-weighted) in the right ear. For masker speech of longer duration, the RMS normalization also included a gated loudness-based procedure per EBU R128 Standard (Rodriguez, 2015). Speech materials convolved with KEMAR HRTFs recorded at $\pm 90^\circ$ carried the naturally occurring ILDs and ILDs associated with different virtual locations.

The use of VAS and non-individual HRTFs to measure binaural intelligibility benefits on children was motivated by two practical considerations. First, the headphone-based reproduction of VAS to create a monaural listening condition provided greater attenuation than the 20–50 dB from blocked ear methods in free-field testing (Irving & Moore, 2011), particularly in children with smaller and more variable ear canal size (Fels, 2008; Stinson & Lawton, 1989). When presented over headphones, a monaural condition is effectively created by digitally muting the sounds in one ear, simulating a perfectly occluded ear. Hearing through bone conduction in the simulated occluded ear should be minimal through the circumaural headphones with cushioned pad used in this study. Second, even though growing head size may alter the physiological range of ITD and ILD for each child (Clifton et al., 1988; Fels et al., 2004), extreme interaural cues available from the $\pm 90^\circ$ azimuthal locations from KEMAR likely have magnitudes greater than those from most children’s own HRTFs. In this study, we measured children’s ability to utilize interaural cues in SRM by giving each participant access to the same set of auditory cues from KEMAR HRTFs, rather than quantifying access to ITDs/ILDs through their own HRTFs.

Experimental Setup and Procedure

Participants were seated in a double-walled sound booth (Acoustic Systems, Austin, TX) during testing. A monitor screen with a computer mouse was provided for participants to initiate sentence playback for each trial. Auditory stimuli were presented through a pair of Sennheiser HD600 (Wedemark, Germany) circumaural headphones that were connected to an RME Babyface Pro DSP sound card. Participants were instructed to listen for the target sentences and verbally repeat back all the words heard. An examiner sitting outside the sound booth listened for the verbal responses. For scoring, the sentence for each trial was

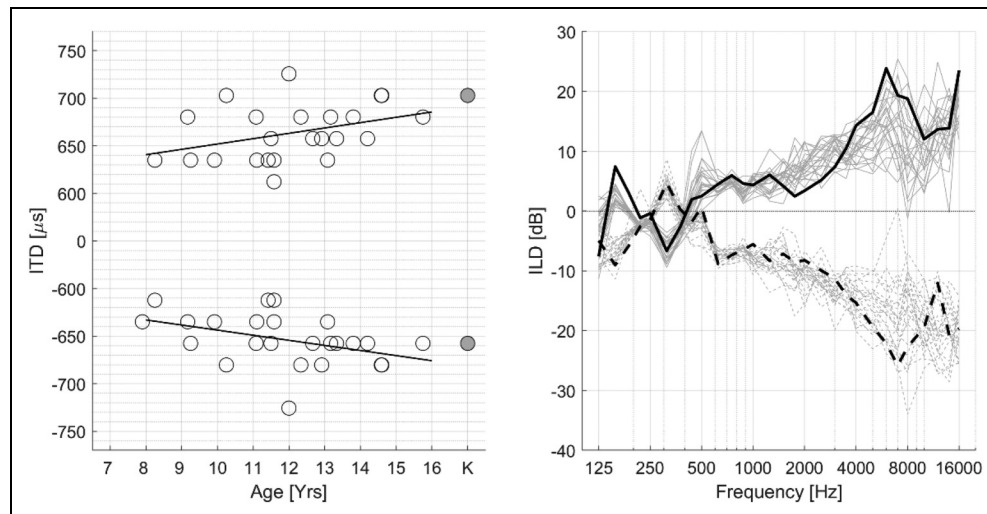


Figure 2. ITD and ILD from individual measurements from 24 children. Measurements from KEMAR manikin were included on the same plot in the gray symbol (“K”) or dark lines (ILD plot). ITDs and ILDs were calculated from two source positions at $\pm 90^\circ$ azimuth. Negative values in ITD and ILD denote locations in the left hemifield. HRTFs were low-pass filtered at 1 kHz for ITD calculation. Abbreviations: ITD = interaural time difference; ILD = interaural level difference.

displayed to the examiner with the keywords highlighted and accompanied by a checkbox. The examiner checked all the keywords were correctly repeated for each trial. Several undergraduate and graduate students served as the examiner during data collection; many of them were native American English speakers. The impact of non-native English-speaking examiners on the scoring was minimal due to the software interface for collecting responses.

For each participant, the quiet condition was always presented first. This provided listeners an opportunity to practice listening for the target voice. The noise conditions with multiple experimental runs were presented in pseudo-randomized order. All adult listeners and most children were tested for three runs in each condition with maskers; due to time constraints, some of the children were repeated in two runs for each condition with maskers. All adults completed the study in a single session of 2 h; all child participants completed the full test protocol in 1–2 sessions of up to 2 h each with breaks. On each test trial, the masker speech always started 2 s before the target sentence.

An adaptive one-down-one-up procedure (Levitt, 1971) was used to estimate SRT at 50% accuracy for each test run. A sentence was considered correct if two or more keywords were correctly identified. The initial presentation level of the target speech was set at 60 dB SPL (re 20 μ Pa). During a test run, the target intensity varied based on the listener’s response, that is, increasing after an incorrect sentence response and decreasing after a correct response. In conditions with babble noise, the maskers were always fixed at 55 dBA SPL (re 20 μ Pa). The initial step size was 8 dB until the first reversal, after which the step size reduced to 4 dB and subsequently 2 dB after the second reversal. Each experimental run was terminated after seven reversals. Three pre-determined

randomized sentence lists were created, such that each list contained the same subset of 597 AuSTIN sentences with three keywords but differed in order. For each participant, one of the lists was selected, and the sentences were assigned to each trial in the order arranged in the list. Each sentence was presented once without repetition within or across trials or test runs.

Twenty-four children also participated in an additional 1-h session to measure individual HRTFs and head size. These measurements were correlated with behavioral performances to understand how children’s access to natural binaural cues from their own HRTFs affects their usage of such cues from KEMAR HRTFs. Figure 2 shows the ITDs and ILDs calculated from individual HRTFs from these 24 children. ITD was calculated by first low-pass filtering the HRTFs at 1 kHz, then cross-correlating the binaural signals to find the interaural delay at maximum correlation. The frequency-dependent ILDs were derived by comparing the intensity between interaural HRTFs in each frequency bin after Fourier Transformation. Spearman’s correlation showed that ITDs are correlated with age ($r[23] = .46$, $p = .029$). Between 6 and 15 years old, children gain an averaged ~ 50 μ s in ITD.

Defining Intelligibility Benefits From Individual Auditory Cues

In this study, SRT was defined as the target speech level yielding a 50% keyword accuracy in quiet, and the SNR at which 50% accuracy was achieved in conditions with masker. SRTs were estimated by extracting the adaptive tracking data and using the maximum likelihood estimation

(MLE) method described by Frund et al. (2011). The MLE method is more robust than reversal-averaging in reducing artifacts from inattention in threshold estimation among children (Litovsky, 2005) who may be at higher risk of fatigue toward the end of a test run. To estimate SRT in each test run, all trial responses were combined and converted to percent correctly repeated keywords for each tested SNR or target SPL. A sigmoid curve was fitted to all trial data, with SRT derived at the 50% accuracy point on the curve. The psychometric function fit resulted in an averaged R^2 of 0.85. For individual curve fit $R^2 < 0.9$, the estimated SRT deviated <1 dB from a repeated run of the same test condition. For each test condition, an averaged SRT was calculated across repeated experimental runs.

By comparing SRTs measured across the four conditions with masker, we calculated the intelligibility benefits from individual and multiple auditory cues using the following equations:

1. Total Unmasking Benefits = $SRT_{Co-located, Monaural} - SRT_{Separated, Binaural}$
2. SRM = $SRT_{Co-located, Binaural} - SRT_{Separated, Binaural}$
3. Binaural Squelch = $SRT_{Separated, Monaural} - SRT_{Separated, Binaural}$
4. Head Shadow = $SRT_{Co-located, Monaural} - SRT_{Separated, Monaural}$
5. Binaural redundancy = $SRT_{Co-located, Monaural} - SRT_{Co-located, Binaural}$
6. Interaural Differences = $SRM - Head\ Shadow = Binaural\ Squelch - Binaural\ redundancy = (SRT_{Co-located, Binaural} - SRT_{Separated, Binaural}) - (SRT_{Co-located, Monaural} - SRT_{Separated, Monaural})$

Results

All statistical analyses were conducted in R (Version 4.0.1) and RStudio, with the “lme4” package for linear mixed-effects modeling and the “nlme” package for linear regression fitting. The a priori level of significance was set at $\alpha = .05$. As a general approach to examining the age effect, chronological age in years was modeled as a continuous variable in the linear regression and linear mixed-effects models fitted to the child data. To compare children’s data with adults, a two-sample one-sided t -test was used to evaluate if predicted SRTs or intelligibility benefits at 15 years old, end of the age range tested, were poorer than those of adults.

Speech Intelligibility

Figure 3 shows the SRTs as a function of age in years for the quiet condition, and all four conditions with maskers. SRTs for sentence recognition in quiet were reported as the target dBA SPL and did not improve over the age range among the children tested (Figure 3a). For all conditions with maskers, the age effect on SRTs was better described by a linear regression rather than a non-linear limited growth function (e.g., power function). Hence, a linear mixed-effects model was fitted to the SRTs with age, spatial separation,

and ear condition as the fixed effects and individual children as the random effect. The initial model contained only the main effects; the addition of interaction between spatial separation and ear condition significantly improved the model prediction ($\chi^2 = 33.0, p < .0001$). Additional interactions involving age did not further improve the model prediction. The final model contained the significant main effect of age ($F[1,31] = 10.5, p = .003$), spatial separation ($F[1,93] = 996.3, p < .0001$), ear condition ($F[1,93] = 69.3, p < .0001$), and the two-way interaction between spatial separation and ear condition ($F[1,93] = 39.6, p < .0001$).

The significant two-way interaction between spatial separation and ear condition was further examined through post hoc analysis by a linear regression fitted to the SRTs in each condition with maskers. As shown in Figure 3b, there was a significant developmental improvement of SRT over the age range in all conditions with maskers, all $p < .05$. The negative correlations between age in years and SRTs were between $-.37$ and $-.46$ in Pearson’s correlation across conditions with maskers. From the slopes of individual regressions, children on average received an SRT improvement between 3.8 and 5.0 dB over the ~ 10 -year age span.

The predicted SRT for 15 year-olds was significantly higher (all $p < .05$) than the averaged SRT from adults by ~ 1.3 dB in both monaural listening conditions and up to 3.8 dB in the binaural separated condition. On average, children had not reached adult-like performance by age 15 years when recalling sentences in a two-talker babble masker, regardless of the availability of auditory cues for unmasking.

Listening with both ears and with a spatial separation between the target and masker provided intelligibility benefits measured as SRT improvement. In the regressions that predict children’s SRTs, the slopes had overlapping 95% confidence intervals and hence did not differ from each other. The intelligibility benefits were observed directly in the change in intercepts between conditions. When compared to monaural listening, binaural listening provided up to 3.4 dB SRT improvement when the target and masker were spatially separated, but the improvement was <1 dB with co-located target-masker. In contrast, spatial separation provided generally larger SRT improvements of 7.0 and 10.1 dB for monaural and binaural listening, respectively.

Speech Intelligibility Benefits From Auditory Spatial Cues

Speech intelligibility benefits were calculated for each listener using the equations outlined above. Figure 4 illustrates individual data of the cumulative benefits from each auditory cue: binaural redundancy, head shadow, and interaural differences. For both groups, there were listeners who showed an intelligibility detriment, that is, increased SRT when an auditory cue became available as compared to when it was absent.

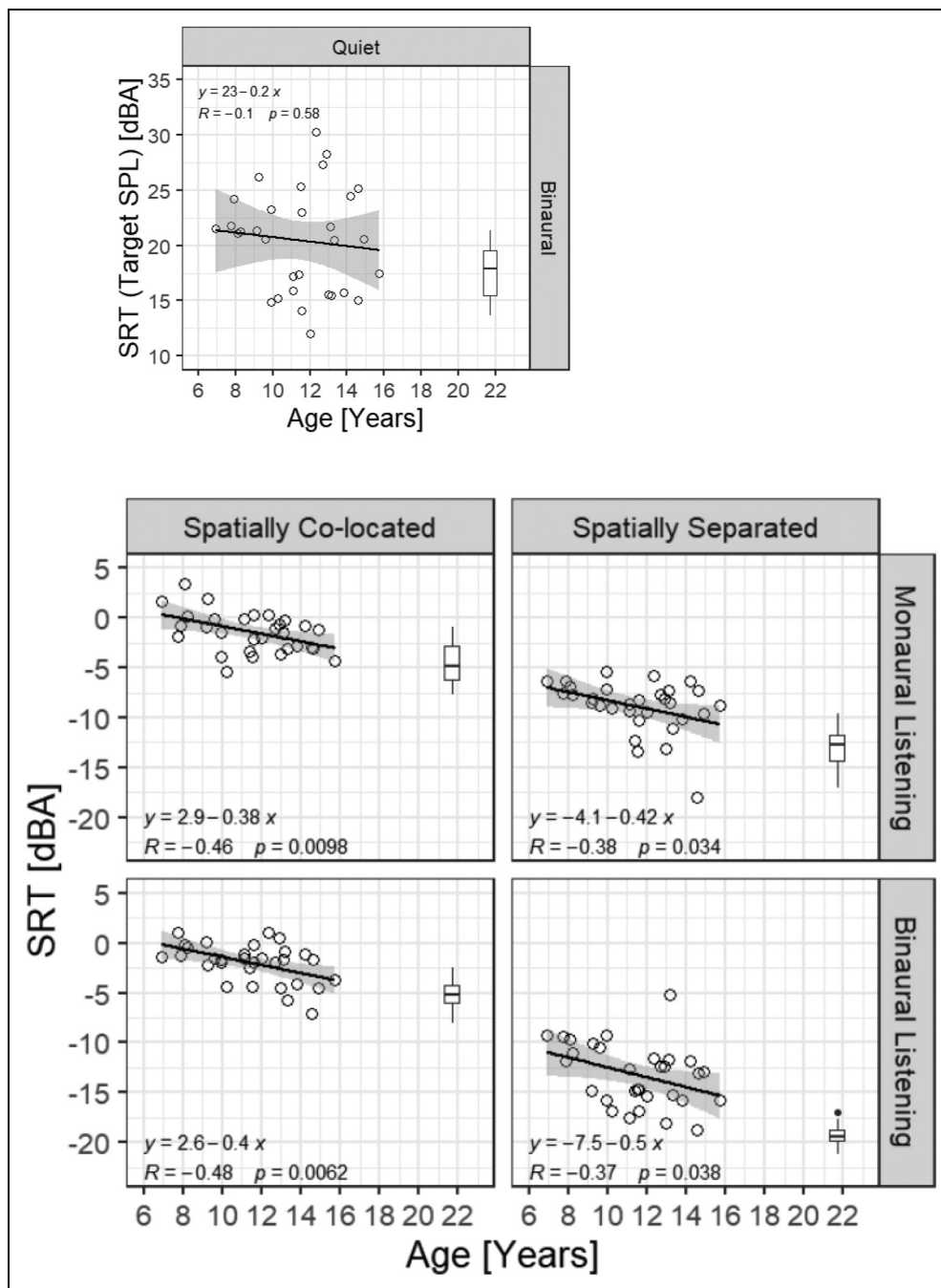


Figure 3. Speech reception thresholds (SRT) as a function of age for children (open circle) in quiet (a) and in four noise conditions (b). Linear regression is fitted to the child data in solid lines, with 95% confidence intervals (the shaded area around a curve). SRTs from adults are reported as a group in a boxplot.

Intelligibility detriments were observed for binaural redundancy (three children and one adult) and interaural difference cues (three children).

The age effect in intelligibility benefits was again better explained by a linear relationship. A mixed effects model was fitted to the intelligibility benefits with age and cue type (i.e., head shadow, binaural redundancy, and interaural differences) as the fixed effects and individual

children as the random effect. The initial model contained only the main effects; the addition of two-way interaction between age and cue type did not improve the model prediction ($\chi^2 = .89, p = .64$). In the final model, there was a significant main effect of cue type ($F[2,123] = 59.9, p < .0001$), but not for age ($F[1,123] = 2.0, p = .16$). Post hoc analysis through a pairwise *t*-test with Bonferroni corrections was used to compare the effect size of

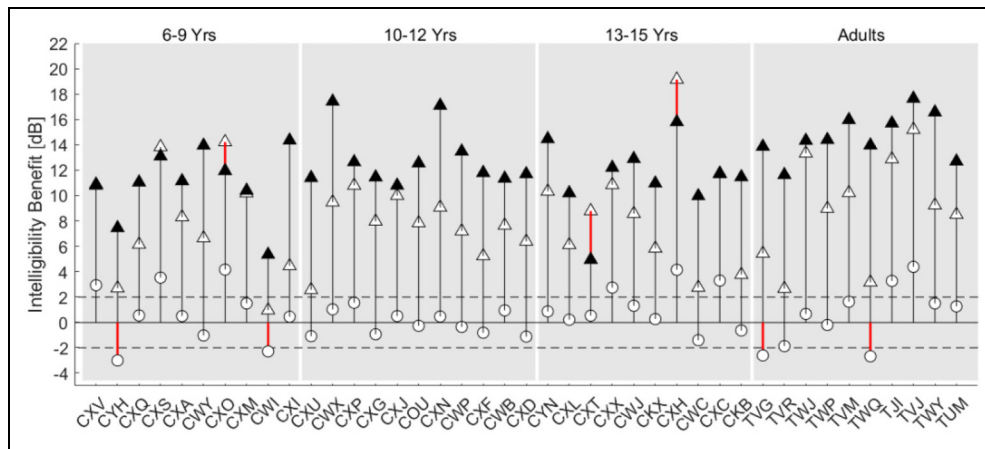


Figure 4. Individual data showing cumulative intelligibility benefit received from each auditory cue: binaural redundancy (dB increase from 0 to circle), head shadow (dB increase from circle to open triangle), and interaural differences (dB increase from open triangle to closed triangle). Listeners are ordered by chronological age and arranged in four age groups. Intelligibility detriments (dB reduction >2 dB from a specific cue) are identified by red vertical lines.

intelligibility benefits from each auditory cue; all pair combinations were statistically significant, all p 's < .0001. Intelligibility benefit was the largest from monaural head shadow ($M = 7.5$ dB, $SD = 2.4$), followed by interaural differences ($M = 4.2$ dB, $SD = 3.6$), and binaural redundancy ($M = 0.6$ dB, $SD = 1.9$).

There was a weak and non-significant developmental improvement of intelligibility benefits from individual cues. Figure 5 shows intelligibility benefits as a function of age in years from individual and combinations of auditory cues, each with a linear regression fitted to the child data. From the regressions, the slopes were not statistically significant even though all positive. On average, the intelligibility benefits increased by no more than 1 dB between 6 and 15 years of age.

When compared with adults, on average, children demonstrated a significantly smaller value of SRM ($t[9] = 5.2$, $p = .0002$), binaural squelch ($t[9] = 2.7$, $p = .012$), and total unmasking benefit ($t[9] = 5.3$, $p = .0003$). The adult-child differences were not significant for benefits from head shadow ($t[9] = 1.2$, $p = .14$) or binaural redundancy ($t[9] = -.17$, $p = .56$), and there was a trend toward significance for interaural differences ($t[9] = 1.75$, $p = .057$). On average, adults demonstrated benefits that were 2.9 dB greater than benefits observed in children when all three auditory cues were available (i.e., total unmasking benefits), of which 0.9 dB was due to monaural head shadow and 2.0 dB from interaural differences.

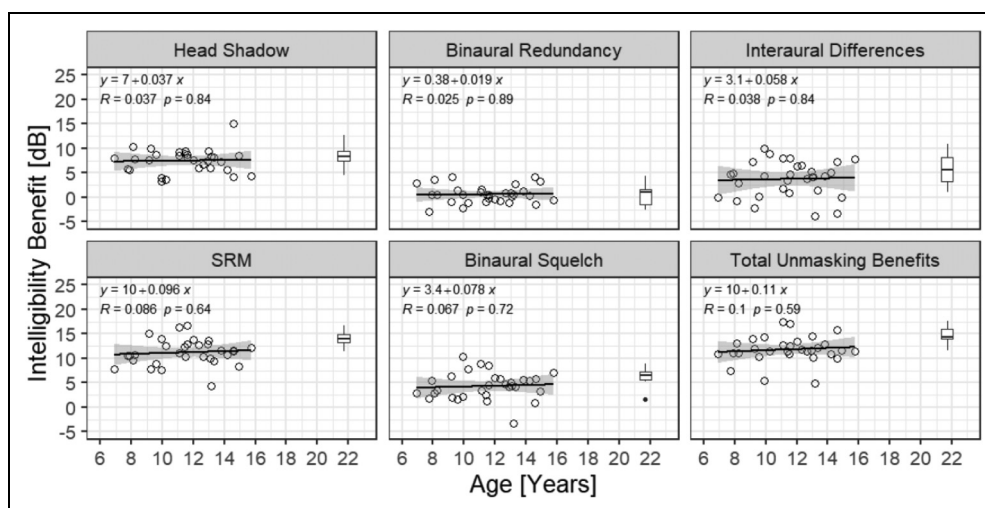


Figure 5. Intelligibility benefits as a function of age from individuals and combinations of auditory cues for unmasking. Linear regression is fitted to the child data in solid lines, with 95% confidence intervals (the shaded area around a curve). Adult data are presented as group data in the boxplot. SRM = spatial release from masking.

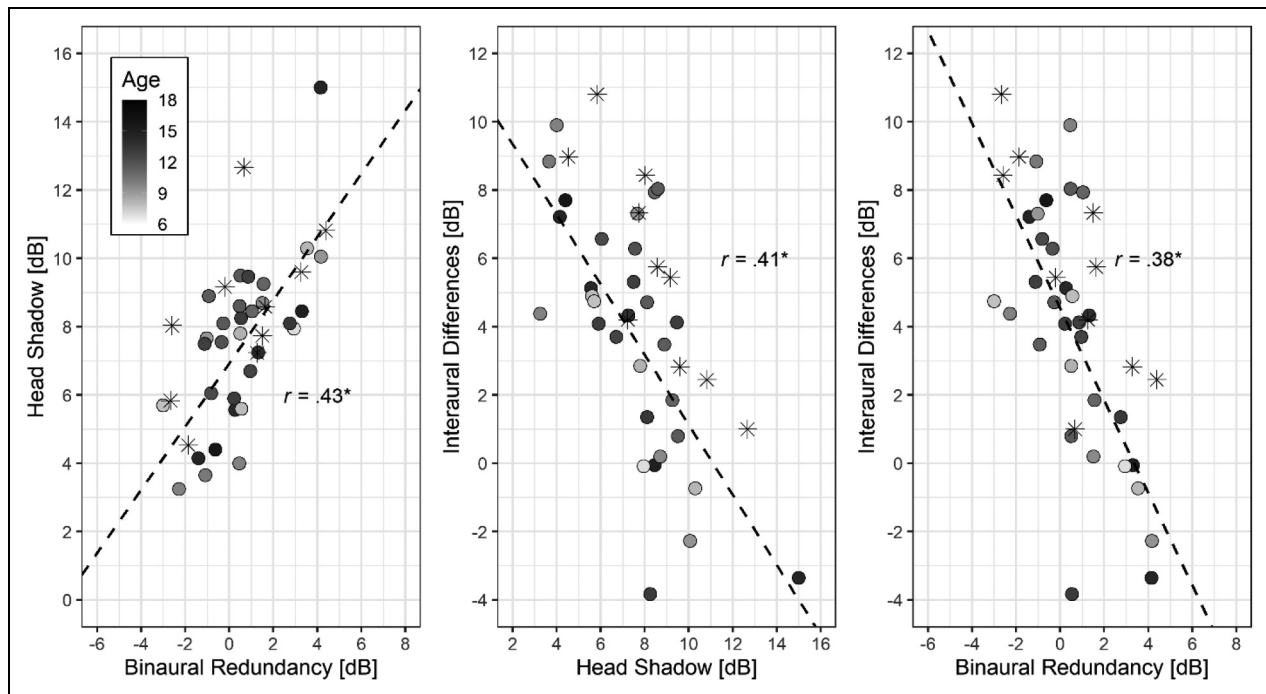


Figure 6. Relations between intelligibility benefit from individual cues in children. Partial correlation is shown in each panel between the pair of cues by controlling the third cue. Data from adults is also plotted but not entered into correlation or regression analyses.

Relationships Between the Use of Individual Cues in Children

Equations (1) to (6) were used to compute the contribution of each auditory cue to the total unmasking benefit from spatial hearing in a mutually exclusive manner. We explored the interrelationships between benefits from individual cues (see Figure 6). Pearson's partial correlations were performed between pairs of individual cues by controlling for the third cue included in Figure 6. Results suggested significant and strong correlations between each pair of auditory cues. To relate all three auditory cues, a linear regression was fitted to intelligibility gains from interaural differences, using benefits from head shadow and binaural redundancy as two independent predictor variables. The omnibus linear regression was significant ($F[2,28] = 16.6, p < .0001, R^2 = .54$, intercept = 8.9). As shown in Figure 6, interaural differences were significantly and negatively predicted by head shadow ($b = -.63, p = .025$) and binaural redundancy ($b = -.77, p = .039$), while the two predictors were correlated positively ($b = .93, p < .0001$). Collinearity between binaural redundancy and head shadow was not of concern with tolerance calculated at 0.52.

Discussion

In this study, we assessed the contribution of auditory cues (i.e., head shadow, binaural redundancy, and interaural differences) to binaural intelligibility benefits in children

between 6 and 15 years of age. When listening in VAS over headphones, children's SRTs in two-talker same-sex babble noise were measured in four test conditions from two spatial configurations (target-masker co-located vs. spatially separated) and two ear conditions (monaural vs. binaural listening). Between 6 and 15 years of age, we observed developmental improvement of SRTs into late childhood in all conditions with a masker. By 15 years old, children had not reached adult-like SRTs regardless of the availability of auditory cues to gain intelligibility benefits. By presenting stimuli in VAS rather than free-field, we were able to better simulate monaural listening and tease apart the contribution of each auditory cue to binaural intelligibility benefits. When the intelligibility benefits from individual auditory cues were examined in isolation, we did not observe a gradual improvement of benefits over the age range tested. However, compared with adults, children at 15 years old demonstrate smaller intelligibility benefits from combinations of these cues, which is likely driven by the ability to benefit from interaural differences.

VAS Using KEMAR HRTFs for Children

Except for those using the LiSN-S test (Cameron & Dillon, 2007), most studies measuring SRM in children have been conducted using loudspeakers in a free-field setup, in which children access auditory spatial cues from their own HRTFs. Both the size of auditory cues and how well children utilize these cues are factors that can influence the amount of

SRM received. By 5–6 months of age, NH infants already have access to approximately 400 μ s ITD from a sound source at $+90^\circ$ (Clifton et al., 1988). For most children, their head circumference will not reach the final adult size until around 7–9 years old, with an estimated maximum ITD close to 600–650 μ s (Fels et al., 2004). Our data of the children's individual HRTFs showed a small increase of around 50 μ s ITD between 6 and 15 years of age, $\sim 10\%$ of adult ITD. By presenting the same set of auditory spatial cues from KEMAR HRTFs to all children, we precluded individual differences in SRM due to variability in children's own HRTFs and were able to specifically assess their ability to use the same set of such cues. The use of headphone presentation also ensured that auditory cues were not altered should children move their heads during stimulus presentation.

A question of interest is whether children can use mismatched ITDs/ILDs (i.e., from KEMAR HRTFs) for spatial hearing in general. For adults, auditory stimuli presented in VAS and over headphones are often perceived as inside the head (Hartmann & Wittenberg, 1996). For children, the extent to which sounds in virtual locations are internalized has not been examined. During the test session, we asked some of the children to provide informal illustrations to indicate where they heard the target and maskers. They were provided a sketch of the head and three options to draw sounds located inside the head, outside the head, or at the ears. Twenty-four children provided these illustrations, of which 23 indicated that sounds were always located outside the head, and one child indicated the sounds at the ears. Surprisingly, the internalization of auditory images was not observed. It is possible that non-zero and exaggerated ITDs/ILDs from the KEMAR HRTFs (cf. their own, Figure 2) might have provided advantages to externalize virtual sounds for children (see review by Best et al., 2020).

Further, we found no influence of individual ITDs/ILDs on children's SRTs or intelligibility benefits. We performed additional stepwise regressions that included ITD, ILD_{HP} , and ILD_{rms} as independent variables to predict SRTs in noise and intelligibility benefits from various cues, all $p > .05$. The ILD_{HP} was calculated by comparing the frequency-dependent HRTFs from two ears and averaged across frequencies ranging from 1.5 to 8 kHz. The ILD_{rms} was calculated by taking the difference in RMS levels of the HRTFs between ears, which better represented the head shadow cues available. Children's access to interaural difference cues from their own HRTFs did not affect their ability to utilize such cues from KEMAR in VAS in this study. Our finding further supports the idea that children can sufficiently use mismatched auditory cues from KEMAR HRTFs for SRM (Cameron et al., 2011; Cameron & Dillon, 2007). This finding is particularly encouraging for using VAS simulated with adult-size, non-individual HRTFs to measure functional spatial hearing outcomes in school-aged children, especially when the spatial separation is sufficiently large.

Access to Interaural Differences Cues

Psychoacoustic literature provides evidence to support that children have access to interaural differences early on in a variety of spatial hearing tasks. In free-field listening, young children's minimum audible angle, the smallest angular separation detectable between two sounds is adult-like at $\sim 1^\circ$ by 4–5 years of age (Litovsky, 1997). By 8–10 years old, children show matured sensitivity to ITD and ILD with adult-like just-noticeable thresholds (Ehlers et al., 2016). When listening to a target signal in noise, children demonstrate improvement in detection thresholds when interaural differences are introduced in the target as compared to when listening diotically (Hall & Grose, 1990; Nozza et al., 1988; Todd et al., 2016; Van Deun et al., 2009). Such threshold improvement is known as the binaural masking level difference (BMLD). Children reach adult-like BMLD by 5 years of age when the target is masked by broadband noise, but demonstrate persistent immaturity with narrow-band noise maskers that produce a greater amount of informational masking (Grose et al., 1997). Using a same-sex two-talker masker with substantial informational masking, Corbin et al. (2017) argued that the poorer ability to use interaural differences may underlie the SRM immaturity observed in children between 8 and 10 years old. In the present study, our data suggest that the difficulty in using interaural differences cues may in fact extend into adolescence and continues to limit unmasking when multiple auditory cues are available.

Age Effect on SRT

In the present study, children demonstrated SRT improvement over time between 6 and 15 years of age in all conditions with maskers. The finding of improving SRTs with age is consistent with results shown by studies using speech maskers with a similar age span (Brown et al., 2010; Buss et al., 2017; Cameron et al., 2011; Corbin et al., 2016; Griffin et al., 2019). Several studies have suggested that the developmental time-course of SRT is better described by the logarithmic transformation of age in years, modeling after a limited growth function with more rapid SRT improvement among younger children than those who are older (Buss et al., 2017; Corbin et al., 2016; Leibold et al., 2020). Our data do not support the limited growth of SRT improvement between the age of 6 and 15 years, as is also shown by Griffin et al. (2019), who studied children in a similar age range. The studies by Buss et al. (2017) and Corbin et al. (2016) included a small number of 5-year-old children. So, one possible explanation is that there may be a window before 6 years of age when children experience a rapid improvement of SRT, after which the growth rate slows down. There is also support from Cameron and Dillon (2007), who analyzed the developmental time-course by grouping the children by year. They showed that the annual SRT improvement between 5 and 6 years of age

was larger than those from each subsequent year as children increase in age.

Our finding on the adult-child difference in SRTs aligns with the general idea that there is a protracted developmental trajectory to understand speech, particularly sentences, when it is masked by speech babble (Brown et al., 2010; Buss et al., 2017; Corbin et al., 2016, 2017; Griffin et al., 2019; Leibold et al., 2020). However, our data show SRT maturation beyond 15 years of age in all noise conditions, which is later than the 9–13 years previously reported in those studies. The protracted development observed in the present study may be the result of overall higher informational masking from the masker speech. When compared to maskers of two different talkers (Brown et al., 2010; Corbin et al., 2017), our approach of using a single same-sex masker speaking two different stories might lead to the increased difficulty to attend to the target because of reduced access to the voice cues from a second masker.

Griffin et al. (2019) reported a linear age effect on SRT with a slope of -0.78 dB/year in the co-located and -0.80 dB/year in the asymmetrically separated condition. When comparing the SRT improvement over the age of 6–12 years between the two studies, children in the present study had generally poorer SRTs in the co-located condition (0 to -2 dB as compared to -2 to -8 dB reported in Griffin et al., 2019), but slightly better SRTs with spatial separation (-11 to -13 dB as compared to -7 to -13 dB in Griffin et al., 2019). One possible explanation is that children may experience smaller binaural redundancy with the $+90^\circ$ azimuthal location than a frontal target. However, by comparing SRTs with frontally co-located target and maskers between the plugged and unplugged ear conditions, both Van Deun et al. (2010) and Corbin et al. (2017) reported a <1 dB benefit from binaural redundancy, similar to the value reported in this study. The better co-located SRTs seen in Griffin et al. (2019) may also be attributable to the child babble masker, possibly with less informational masking than from the same-sex adult masker used in this investigation. The better SRTs in the spatially separated condition suggest that the 180° separation may be advantageous to younger children, but the benefit from the enlarged spatial separation diminishes around 12 years of age.

Age Effect on SRM

SRM is generally thought to emerge before 2–3 years of age (Hess et al., 2018), and its maturation time-course varies between studies using different target and masker materials. Studies using monosyllabic or disyllabic words as the target speech suggested earlier SRM maturation by 5 years of age (Misurelli & Litovsky, 2012, 2015; Murphy et al., 2011; Van Deun et al., 2010), while several other studies used open-set sentences and showed later SRM maturation at 9 years or older (Brown et al., 2010; Cameron et al., 2011; Corbin et al., 2017; Vaillancourt et al., 2008).

Increasing informational masking by using same-sex two-talker babble as the masker also seems to result in with more protracted SRM maturation (Brown et al., 2010; Cameron et al., 2011; Corbin et al., 2017) than using similar speech maskers of babble from children (Griffin et al., 2019).

In the present study, we used open-set sentences for the target speech and a same-sex two-talker babble as the masker, with the aim to maximize informational masking and promote the use of auditory spatial cues for SRM in children. In contrary to existing literature, we observed SRM immaturity into adolescence (i.e., 15 years of age). When all three auditory cues were available, the adult-child difference of ~ 3 dB SRM was mainly driven by children's poorer ability to use interaural differences. It is unclear whether the prolonged SRM (cf. interaural differences) maturation is related to the 180° fixed angular separation in the present study, which is much larger than the 60° – 90° separation used in previous work. But our finding on the small adult-child difference in SRM from head shadow (<1 dB) is consistent with past work that children reach maturation early on using this cue for SRM (Corbin et al., 2017; Griffin et al., 2019; Misurelli & Litovsky, 2015).

The Role of Auditory Cues in Spatial Release From Masking

Several studies show an averaged 2–5 dB SRM from monaural head shadow in both children and adults by either directly measuring it with a blocked ear method or comparing between symmetrical and asymmetrical SRMs (Cameron et al., 2011; Corbin et al., 2017; Griffin et al., 2019; Misurelli & Litovsky, 2015; Van Deun et al., 2010). When compared with these studies using a 60° – 90° separation, the 180° separation in the present study led to a larger monaural head shadow at 7–8 dB on average. We observed an averaged 3–5 dB intelligibility benefits from interaural differences that were similar to the values derived from comparisons in Corbin et al. (2017), who also used open-set sentences as targets and a same-sex two-talker masker but a 90° separation. Contrary to our prediction, there does not seem to be an increase in intelligibility benefits from interaural differences by increasing the spatial separation beyond 90° .

Note that binaural redundancy in general provided <1 dB in unmasking benefits for both children and adults, much smaller than the 2.6 dB reported by Dieudonné and Francart (2019) using a similar VAS-based method and a 90° fixed angular separation. We speculate that the discrepancy in binaural redundancy effect sizes across the two studies may be related to the auditory cue and the noise masker. First, even though in our study the input from the contralateral (left) ear was coherent, it had a lower intensity by 4.6 dB with a $+90^\circ$ azimuthal co-location of the target and masker due to the acoustic head shadow. The low-

intensity coherent input from the added ear (Figure 1) may simply not be enough to provide substantial binaural redundancy. Second, it is known that amplitude modulation facilitates the calculation of interaural coherence and thus enhances the binaural redundancy effect. The smaller binaural redundancy effect in the present study might therefore be the result of the two-talker masker, which has less temporal modulation than the single-talker masker used in Dieudonné and Francart (2019).

Despite our best effort to isolate each individual auditory cue to assess their contribution to total unmasking benefits, there are several considerations worth discussing when evaluating effect sizes. In a test condition with a substantial amount of informational masking (e.g., co-located monaural listening), the adaptive SRT procedure may include several trials at > 0 dB SNR after the initial reversal. The positive SNRs introduce an additional level cue for release from masking (Swaminathan et al., 2015) that is not available in conditions when the auditory spatial cues are present, in which the SNRs are more consistently < 0 dB from trial to trial. Leibold et al. (2020) investigated a sex-mismatch benefit in a release from masking and identified the spurious level cue as a potential issue in underestimating SRTs for the baseline conditions, which subsequently limits the effect size of the auditory cue of interest. In the present study, the effect size of intelligibility benefits may be underestimated due to the same issue.

Another consideration is that intelligibility benefits from interaural differences were derived rather than directly measured. In Figure 1, the head shadow and binaural redundancy cues were introduced one at a time from the baseline co-located monaural listening condition, by either displacing the maskers (i.e., separated monaural listening) or adding the second ear (i.e., co-located binaural listening). But the interaural differences cue was always introduced along with another cue; its contribution to SRM was derived by subtracting the other cue's measured benefit from the combined benefits. We took the similar assumption as Dieudonné and Francart (2019) that each auditory cue's contribution to unmasking is independent, because these cues do not have shared mechanisms for unmasking. This assumption may not completely reflect children's use of these cues when all are available, particularly as we observed the trade-off between the benefit from interaural differences versus the benefit from head shadow. It warrants future investigation to understand the relative independence of benefits drawn from individual cues where the interaural differences cue can be introduced in isolation.

Conclusions

This study examined the contribution of individual auditory cues to binaural speech intelligibility benefits in NH children over a wide age range between 6 and 15 years old. The ability to recall everyday sentences in same-sex two-talker babble

noise continued to improve into late adolescence, regardless of the availability of auditory spatial cues. Using a 180° angular separation, the maximum natural spatial separation between target and masker, we quantified the binaural intelligibility benefits from each auditory cue, including head shadow, binaural redundancy, and interaural differences, as well as from combinations of these cues. We observed a prolonged maturation of intelligibility benefits when more than one auditory cue was available, including SRM, binaural squelch, and total unmasking benefits. The adult-child difference in SRM (3.9 dB) seemed to be primarily driven by children's poorer use of interaural differences (2.0 dB) rather than head shadow (0.9 dB).

Acknowledgments

The authors are grateful to all the research participants for their time in this study. The authors thank Alan Kan and Sara Misurelli for helpful conversations during the initial development of the study, Shelly Godar for participant recruitment and audiological assistance, and Molly Beier, Ruth White, Brianna Ralston, and Elisa Carlon for assistance during data collection.

Author's Notes

Portions of this study were presented at the 177th Acoustical Society of America Meeting at Louisville, KY.

Declaration of Conflicting of Interest

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.



Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the National Institute of Health (NIH) National Institute on Deafness and Other Communication Disorders Grants No. R01DC003083 and R01DC008365 (R.Y.L.), and in part by NIH Eunice Kennedy Shriver National Institute of Child Health and Human Development Grant No. T32HD007489 and U54HD090256 and the University of Wisconsin-Madison.

Data Accessibility Statement

Data will be made available upon reasonable request.

ORCID iDs

Ruth Y. Litovsky  <https://orcid.org/0000-0002-6221-937X>
Z. Ellen Peng  <https://orcid.org/0000-0002-3487-2144>

References

- Best V., Baumgartner R., Lavandier M., Majdak P., & Kopčo N. (2020). Sound externalization: A review of recent research. *Trends in Hearing*, 24, 233121652094839. <https://doi.org/10.1177/2331216520948390>
- Blauert J. (1997). *Spatial hearing: The psychophysics of human sound localization*. MIT Press.

- Bronkhorst A. W. (2015). The cocktail-party problem revisited: Early processing and selection of multi-talker speech. *Attention, Perception, & Psychophysics*, *77*, 1465–1487. <https://doi.org/10.3758/s13414-015-0882-9>
- Bronkhorst A. W., & Plomp R. (1988). The effect of head-induced interaural time and level differences on speech intelligibility in noise. *The Journal of the Acoustical Society of America*, *83*(4), 1508–1516. <https://doi.org/10.1121/1.395906>
- Brown D. K., Cameron S., Martin J. S., Watson C., & Dillon H. (2010). The North American listening in spatialized noise-sentences test (NA LiSN-S): Normative data and test-retest reliability studies for adolescents and young adults. *Journal of the American Academy of Audiology*, *21*(10), 629–641. <https://doi.org/10.3766/jaaa.21.10.3>
- Brunhart D. S., Simpson B. D., Ericson M. A., & Scott K. R. (2001). Informational and energetic masking effects in the perception of multiple simultaneous talkers. *The Journal of the Acoustical Society of America*, *110*(5), 2527–2538. <https://doi.org/10.1121/1.1408946>
- Buss E., Leibold L. J., Porter H. L., & Grose J. H. (2017). Speech recognition in one- and two-talker maskers in school-age children and adults: Development of perceptual masking and glimpsing. *The Journal of the Acoustical Society of America*, *141*(4), 2650–2660. <https://doi.org/10.1121/1.4979936>
- Cameron S., & Dillon H. (2007). Development of the listening in spatialized noise-sentences test (LISN-S). *Ear and Hearing*, *28*(2), 196–211. <https://doi.org/10.1097/AUD.0b013e318031267f>
- Cameron S., Dillon H., & Newall P. (2006). The listening in spatialized noise test: Normative data for children. *International Journal of Audiology*, *45*(2), 99–108. <https://doi.org/10.1080/14992020500377931>
- Cameron S., Glyde H., & Dillon H. (2011). Listening in spatialized noise-sentences test (LiSN-S): Normative and retest reliability data for adolescents and adults up to 60 years of age. *Journal of the American Academy of Audiology*, *22*(10), 697–709. <https://doi.org/10.3766/jaaa.22.10.7>
- Ching T. Y. C., van Wanrooy E., Dillon H., & Carter L. (2011). Spatial release from masking in normal-hearing children and children who use hearing aids. *The Journal of the Acoustical Society of America*, *129*(1), 368–375. <https://doi.org/10.1121/1.3523295>
- Clifton R. K., Gwiazda J., Bauer J. A., Clarkson M. G., & Held R. M. (1988). Growth in head size during infancy: Implications for sound localization. *Developmental Psychology*, *24*(4), 477–483. <https://doi.org/10.1037/0012-1649.24.4.477>
- Corbin N. E., Bonino A. Y., Buss E., & Leibold L. J. (2016). Development of open-set word recognition in children. *Ear & Hearing*, *37*(1), 55–63. <https://doi.org/10.1097/AUD.0000000000000201>
- Corbin N. E., Buss E., & Leibold L. J. (2017). Spatial release from masking in children: Effects of simulated unilateral hearing loss. *Ear & Hearing*, *38*(2), 223–235. <https://doi.org/10.1097/AUD.0000000000000376>
- Dawson P. W., Hersbach A. a., & Swanson B. a. (2013). An adaptive Australian sentence test in noise (Austin). *Ear and Hearing*, *34*(5), 592–600. <https://doi.org/10.1097/AUD.0b013e31828576fb>
- Dieudonné B., & Francart T. (2019). Redundant information is sometimes more beneficial than spatial information to understand speech in noise. *Ear and Hearing*, *40*(3), 545–554. <https://doi.org/10.1097/AUD.0000000000000660>
- Ehlers E., Kan A., Winn M. B., Stoelb C., & Litovsky R. Y. (2016). Binaural hearing in children using Gaussian enveloped and transposed tones. *The Journal of the Acoustical Society of America*, *139*(4), 1724–1733. <https://doi.org/10.1121/1.4945588>
- Epstein M., & Florentine M. (2009). Binaural loudness summation for speech and tones presented via earphones and loudspeakers. *Ear and Hearing*, *30*(2), 234–237. <https://doi.org/10.1097/AUD.0b013e3181976993>
- Fels J., Buthmann P., & Vorländer M. (2004). Head-related transfer functions of children. *Acta Acustica United with Acustica*, *90*(5), 918–927.
- Fels J. (2008). From children to adults: How binaural cues and ear canal impedances grow. Doctoral diss., Acoustical Society of America. <https://doi.org/10.1121/1.3020603>
- Freyman R. L., Balakrishnan U., & Helfer K. S. (2001). Spatial release from informational masking in speech recognition. *The Journal of the Acoustical Society of America*, *109*(5), 2112–2122. <https://doi.org/10.1121/1.1354984>
- Frund I., Haenel N. V., & Wichmann F. A. (2011). Inference for psychometric functions in the presence of nonstationary behavior. *Journal of Vision*, *11*(6), 16–16. <https://doi.org/10.1167/11.6.16>
- Garadat S. N., & Litovsky R. Y. (2007). Speech intelligibility in free field: Spatial unmasking in preschool children. *The Journal of the Acoustical Society of America*, *121*(2), 1047–1055. <https://doi.org/10.1121/1.2409863>
- Griffin A. M., Poissant S. F., & Freyman R. L. (2019). Speech-in-noise and quality-of-life measures in school-aged children With normal hearing and With unilateral hearing loss. *Ear & Hearing*, *40*(4), 887–904. <https://doi.org/10.1097/AUD.0000000000000667>
- Grose J. H., Hall III J. W., & Dev M. B. (1997). MLD in children: Effects of signal and masker bandwidths. *Journal of Speech, Language, and Hearing Research*, *40*(4), 955–959. <https://doi.org/10.1044/jslhr.4004.955>
- Hall J. W., & Grose J. H. (1990). The masking-level difference in children. *Journal of the American Academy of Audiology*, *1*(2), 81–88.
- Hartmann W. M., & Wittenberg A. (1996). On the externalization of sound images. *The Journal of the Acoustical Society of America*, *99*(6), 3678–3688. <https://doi.org/10.1121/1.414965>
- Hawley M. L., Litovsky R. Y., & Culling J. F. (2004). The benefit of binaural hearing in a cocktail party: Effect of location and type of interferer. *Journal of the Acoustical Society of America*, *115*(2), 833–843. <https://doi.org/10.1121/1.1639908>
- Hess C. L., Misurelli S. M., & Litovsky R. Y. (2018). Spatial release from masking in 2-year-olds with normal hearing and with bilateral cochlear implants. *Trends in Hearing*, *22*, 233121651877556. <https://doi.org/10.1177/2331216518775567>
- Ihlefeld A., & Shinn-Cunningham B. (2008). Spatial release from energetic and informational masking in a selective speech identification task. *The Journal of the Acoustical Society of America*, *123*(6), 4369–4379. <https://doi.org/10.1121/1.2904826>
- Irving S., & Moore D. R. (2011). Training sound localization in normal hearing listeners with and without a unilateral ear plug. *Hearing Research*, *280*(1–2), 100–108. <https://doi.org/10.1016/j.heares.2011.04.020>
- Johnstone P. M., & Litovsky R. Y. (2006). Effect of masker type and age on speech intelligibility and spatial release from masking in children and adults. *The Journal of the Acoustical*

- Society of America*, 120(4), 2177–2189. <https://doi.org/10.1121/1.2225416>
- Jones G. L., & Litovsky R. Y. (2008). Role of masker predictability in the cocktail party problem. *The Journal of the Acoustical Society of America*, 124(6), 3818–3830. <https://doi.org/10.1121/1.2996336>
- Kidd G., Mason C. R., Swaminathan J., Roverud E., Clayton K. K., & Best V. (2016). Determining the energetic and informational components of speech-on-speech masking. *The Journal of the Acoustical Society of America*, 140(1), 132–144. <https://doi.org/10.1121/1.4954748>
- Leibold L. J., Browning J. M., & Buss E. (2020). Masking release for speech-in-speech recognition due to a target/masker Sex mismatch in children with hearing loss. *Ear and Hearing*, 41(2), 259–267. <https://doi.org/10.1097/AUD.0000000000000752>
- Leibold L. J., Buss E., & Calandruccio L. (2018). Developmental effects in masking release for speech-in-speech perception due to a target/masker sex mismatch. *Ear and Hearing*, 39(5), 935–945. <https://doi.org/10.1097/AUD.0000000000000554>
- Levitt H. (1971). Transformed up-down methods in psychoacoustics. *The Journal of the Acoustical Society of America*, 49(2B), 467–477. <https://doi.org/10.1121/1.1912375>
- Litovsky R. Y. (1997). Developmental changes in the precedence effect: Estimates of minimum audible angle. *The Journal of the Acoustical Society of America*, 102(3), 1739–1745. <https://doi.org/10.1121/1.420106>
- Litovsky R. Y. (2005). Speech intelligibility and spatial release from masking in young children. *The Journal of the Acoustical Society of America*, 117(5), 3091–3099. <https://doi.org/10.1121/1.1873913>
- Misurelli S. M., & Litovsky R. Y. (2012). Spatial release from masking in children with normal hearing and with bilateral cochlear implants: Effect of interferer asymmetry. *The Journal of the Acoustical Society of America*, 132(1), 380–391. <https://doi.org/10.1121/1.4725760>
- Misurelli S. M., & Litovsky R. Y. (2015). Spatial release from masking in children with bilateral cochlear implants and with normal hearing: Effect of target-interferer similarity. *The Journal of the Acoustical Society of America*, 138(1), 319–331. <https://doi.org/10.1121/1.4922777>
- Murphy J., Summerfield A. Q., O'Donoghue G. M., & Moore D. R. (2011). Spatial hearing of normally hearing and cochlear implanted children. *International Journal of Pediatric Otorhinolaryngology*, 75(4), 489–494. <https://doi.org/10.1016/j.ijporl.2011.01.002>
- Nozza R. J., Wagner E. F., & Crandell M. A. (1988). Binaural release from masking for a speech sound in infants, preschoolers, and adults. *Journal of Speech, Language, and Hearing Research*, 31(2), 212–218. <https://doi.org/10.1044/jshr.3102.212>
- Plomp R. (1976). Binaural and monaural speech intelligibility of connected discourse in reverberation as a function of azimuth of a single competing sound source (speech or noise). *Acustica*, 34(4), 200–211.
- Rodriguez A. (2015). Audio quality in audio description: Getting the priorities right (June 2015). 2015 *IEEE International Symposium on Broadband Multimedia Systems and Broadcasting*, 1–3. <https://doi.org/10.1109/BMSB.2015.7177235>.
- Stinson M. R., & Lawton B. W. (1989). Specification of the geometry of the human ear canal for the prediction of sound-pressure level distribution. *The Journal of the Acoustical Society of America*, 85(6), 2492–2503. <https://doi.org/10.1121/1.397744>
- Swaminathan J., Mason C. R., Streeter T. M., Best V., Kidd G., & Patel A. D. (2015). Musical training, individual differences and the cocktail party problem. *Scientific Reports*, 5(1), 1–11. <https://doi.org/10.1038/srep11628>
- TIME USA (n.d.). TIME for Kids. Retrieved November 6, 2020, from <https://www.timeforkids.com/>.
- Todd A. E., Goupell M. J., & Litovsky R. Y. (2016). Binaural release from masking with single- and multi-electrode stimulation in children with cochlear implants. *Journal of the Acoustical Society of America*, 140(1), 59–73. <https://doi.org/10.1121/1.4954717>
- Vaillancourt V., Laroche C., Giguère C., & Soli S. D. (2008). Establishment of age-specific normative data for the Canadian French version of the hearing in noise test for children. *Ear and Hearing*, 29(3), 453–466. <https://doi.org/10.1097/01.aud.0000310792.55221.0c>
- Van Deun L., van Wieringen A., Van den Bogaert T., Scherf F., Offeciers F. E., Van de Heyning P. H., & Wouters J. (2009). Sound localization, sound lateralization, and binaural masking level differences in young children with normal hearing. *Ear and Hearing*, 30(2), 178–190. <https://doi.org/10.1097/AUD.0b013e318194256b>
- Van Deun L., van Wieringen A., & Wouters J. (2010). Spatial speech perception benefits in young children with normal hearing and cochlear implants. *Ear and Hearing*, 31(5), 702–713. <https://doi.org/10.1097/AUD.0b013e3181e40dfe>
- Yuen K. C. P., & Yuan M. (2014). Development of spatial release from masking in mandarin-speaking children with normal hearing. *Journal of Speech, Language, and Hearing Research*, 57(5), 2005–2023. https://doi.org/10.1044/2014_JSLHR-H-13-0060nn