



Masked speech recognition by 6–13-year-olds with early-childhood otitis media: effects of acoustic condition and otologic history

Shno Koiek, Christian Brandt, Sören Möller, Harvey Dillon & Tobias Neher

To cite this article: Shno Koiek, Christian Brandt, Sören Möller, Harvey Dillon & Tobias Neher (2025) Masked speech recognition by 6–13-year-olds with early-childhood otitis media: effects of acoustic condition and otologic history, *International Journal of Audiology*, 64:3, 224–231, DOI: [10.1080/14992027.2024.2348506](https://doi.org/10.1080/14992027.2024.2348506)

To link to this article: <https://doi.org/10.1080/14992027.2024.2348506>



© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group on behalf of British Society of Audiology, International Society of Audiology, and Nordic Audiological Society.



Published online: 20 May 2024.



Submit your article to this journal [↗](#)



Article views: 1174



View related articles [↗](#)



View Crossmark data [↗](#)

Masked speech recognition by 6–13-year-olds with early-childhood otitis media: effects of acoustic condition and otologic history

Shno Koiek^{a,b,*} , Christian Brandt^{a,b} , Sören Möller^{a,c} , Harvey Dillon^{d,e}  and Tobias Neher^{a,b} 

^aDepartment of Clinical Research, University of Southern Denmark, Odense, Denmark; ^bResearch Unit for ORL – Head & Neck Surgery and Audiology, Odense University Hospital & University of Southern Denmark, Odense, Denmark; ^cOPEN, Open Patient data Explorative Network, Odense University Hospital, Odense, Denmark; ^dDepartment of Linguistics, Macquarie University, Sydney, Australia; ^eManchester Centre for Audiology and Deafness, The University of Manchester, Manchester, UK

ABSTRACT

Objective: To investigate speech recognition in school-age children with early-childhood otitis media (OM) in conditions with noise or speech maskers with or without interaural differences. To also investigate the effects of three otologic history factors.

Design: Using headphone presentation, speech recognition thresholds (SRTs) were measured with simple sentences. As maskers, stationary speech-shaped noise (SSN) or two-talker running speech (TTS) were used. The stimuli were presented in a monaural and binaural condition (SSN) or a co-located and spatially separated condition (TTS). Based on the available medical records, overall OM duration, OM onset age, and time since the last OM episode were estimated.

Study sample: 6–13-year-olds with a history of recurrent OM ($N=42$) or without any ear diseases ($N=20$) with normal tympanograms and audiograms at the time of testing.

Results: Mixed-model regression analyses that controlled for age showed poorer SRTs for the OM group (Δ -value = 0.84 dB, $p=0.009$). These appeared driven by the spatially separated, binaural, and monaural conditions. The OM group showed large inter-individual differences, which were unrelated to the otologic history factors.

Conclusions: Early-childhood OM can affect speech recognition in different acoustic conditions. The effects of the otologic history warrant further investigation.

ARTICLE HISTORY

Received 5 October 2023
Revised 16 February 2024
Accepted 15 April 2024

KEYWORDS

Speech perception; otitis media; auditory development; binaural hearing



1. Introduction

Otitis media (OM) is one of the most common early-childhood diseases and causes transient, frequently recurrent, and intermittent conductive hearing loss (CHL), especially in the lower frequencies (e.g. Moore, Hartley, and Hogan 2003). Most OM episodes happen during the first three years of life, with the first episode occurring 6–18 months after birth (e.g. Haggard and Hughes 1991). Early childhood is a sensitive period for the development of auditory abilities. It is possible that interruptions to auditory stimulation caused by CHL during auditory development result in long-term auditory processing deficits, with potential consequences for language development and perception (e.g. Whitton and Polley 2011).

The long-term effects of early-childhood OM on masked speech recognition have been the focus of several studies. Zumach et al. (2009) assessed speech recognition in stationary noise in 7-year-olds ($N=55$) with prospectively documented OM history from birth to 24 months of age. For each child, OM severity was determined based on the number of OM episodes (with more episodes corresponding to more severe OM) and laterality of the disease (with bilateral OM counting twice as much as unilateral OM) during that period. The speech and noise

signals were presented from the same direction, that is, without any interaural differences among them. The analyses showed a moderate correlation between OM severity and speech-in-noise performance, with more severe OM being associated with poorer performance. Keogh et al. (2005) assessed speech recognition in stationary noise with three groups of school-age children: (1) children with <4 OM episodes, (2) children with 4–9 OM episodes, and (3) children with >9 OM episodes. The number of OM episodes was determined based on parental reports. The speech and noise signals were presented diotically, that is, without any interaural differences among them. In contrast to Zumach et al. (2009), these authors found similar speech scores for the three groups of participants.

Masked speech recognition is a complex process that, in many situations, relies on binaural hearing abilities for separating target speech from interfering sounds (e.g. Neher, Fogh, and Koiek 2022; Peng and Litovsky 2021). Apart from binaural (or other auditory) abilities, language skills and cognitive function can also influence masked speech recognition (e.g. Dawes and Bishop 2009). Thus, based on speech-in-noise scores such as those collected for the studies summarised above alone, it can be difficult to determine if any observable speech recognition

CONTACT Tobias Neher  tneher@health.sdu.dk  Department of Clinical Research, University of Southern Denmark, Campusvej 55, 5230 Odense M, Denmark
*Current address: Interacoustics A/S, Audiometer Allé 1, Middelfart, Denmark

© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group on behalf of British Society of Audiology, International Society of Audiology, and Nordic Audiological Society.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

deficits are related to binaural impairments. To reach a conclusion about the effects of early-childhood OM on hearing abilities, measures that control for language skills and cognitive function should be utilised. Evaluating the difference in performance between two corresponding test conditions (i.e. calculating a differential test score) is one way of facilitating this (Dillon and Cameron 2021).

The Listening in Spatialised Noise-Sentences test (LiSN-S) is a method for assessing the ability of children to benefit from spatial cues when trying to segregate target speech from competing speech signals (Cameron and Dillon 2007). Using virtual acoustics, the LiSN-S creates a three-dimensional auditory display under headphones, with target speech presented from in front (0° azimuth) and two speech interferers presented from either also in front or the sides ($\pm 90^\circ$ azimuth). In the first condition, the three speech signals are therefore spatially co-located (i.e. without any interaural differences among them), whereas in the other condition they are spatially separated (i.e. with interaural differences among them). When the target speech and speech interferers are spatially separated, spatial release from masking (SRM), which is calculated as the difference between the SRTs measured with and without spatial differences, occurs for normal-hearing listeners.

More recently, a couple of studies investigated the long-term effects of early-childhood OM on SRM. Tomlin and Rance (2014) compared 6–13-year-olds with or without a history of OM in terms of their LiSN-S results. In addition, they examined the influence of OM onset age and overall OM duration on these results. OM history was determined based on parental reports. Relative to the controls, the OM group showed comparable SRTs under co-located conditions but poorer SRTs under spatially separated conditions (Δ -value = ~ 2.4 dB) as well as reduced SRM (Δ -value = ~ 2.2 dB). Moreover, children with an early OM onset age and/or longer overall OM did worse. Following up on these results, Graydon et al. (2017) investigated the long-term effects of early-childhood OM on LiSN-S results in 6–13-year-olds with a documented OM history before age 4. Consistent with Tomlin and Rance (2014), Graydon et al. (2017) observed comparable SRTs under co-located conditions but poorer SRTs under spatially separated conditions (Δ -value = ~ 1.3 dB) as well as reduced SRM (Δ -value = ~ 1.4 dB) for their OM group relative to a control group. They therefore concluded that early-childhood OM can impair the ability to utilise binaural cues for separating target speech from spatially separated speech interferers.

Taken together, there is some evidence that early-childhood OM has negative effects on masked speech recognition in school-age children, and that the progression of the disease (i.e. the number of OM episodes, OM onset age, and overall OM duration) plays a role for these deficits. At a more detailed level, however, there are some inconsistencies. According to two studies, OM-induced speech recognition deficits manifest in situations with spatially separated (but not co-located) speech interferers and thus interaural differences among the competing signals (Graydon et al. 2017; Tomlin and Rance 2014). In contrast, two other studies, which used stationary noise and stimuli without interaural differences, either found an adverse effect of OM based on documented OM records (Zumach et al. 2009) or did not based on parental reports only (Keogh et al. 2005).

The observable inconsistencies could be due to differences in masker type, stimulus presentation, and the way in which OM history was assessed in these studies. To shed more light on these issues, the current study examined speech recognition in school-age children with early-childhood OM in conditions with

noise or speech maskers with or without interaural differences. Using two-talker running speech (TTS) or stationary speech-shaped noise (SSN) as masker, speech recognition thresholds (SRTs) were measured in conditions with and without interaural differences between the competing signals. Furthermore, to assess the ability to benefit from interaural differences, advantage scores were calculated based on the two types of SRTs per masker type. To investigate the effects of OM history on masked speech recognition, three otologic history factors – overall OM duration, age at first OM onset, and the time passed since the last OM episode – were extracted from documented medical records available for the OM group and considered in the statistical analyses.

2. Materials and methods

The current study was evaluated by the Regional Committees on Health Research Ethics for Southern Denmark, and full ethical approval was deemed unnecessary. Therefore, a waiver was granted, as is common practice in such a case. For each participant, written informed consent was obtained from at least one parent. For the OM group, this included permission to obtain access to the child's medical records from the responsible otologist. At the end of the study, all participants received a gift card (corresponding to 120 Danish crowns per visit) for their efforts.

2.1. Participants

Fifty-three children with a documented history of middle-ear infection or effusion (“OM group”) and 22 children without any reported ear diseases (“control group”) were recruited. Eleven children from the OM group and two children from the control group had to be excluded as they either did not pass all inclusion criteria (two children showed type-C tympanograms with static compliance between 0.3 and 1.5 cc and middle-ear pressure less than -100 daPa, and seven children had elevated pure-tone hearing thresholds; see below) or because they decided to withdraw from the study ($N=4$). The remaining 42 children from the OM group (21 female) were aged 6–13 years (mean: 10.1 years; standard deviation, SD: 2.0 years; 25th and 75th percentiles: 8.1 and 11.6 years, respectively). The remaining 20 control children (14 female) were also aged 6–13 years (mean: 10.1 years; SD: 1.9 years; 25th and 75th percentiles: 8.5 and 12.2 years, respectively). An independent *t*-test confirmed no difference in mean age between the two groups ($p > 0.8$).

At the time of testing, all participants fulfilled the following inclusion criteria: (1) type-A tympanogram (static compliance between 0.3 and 1.5 cc and middle-ear pressure between -100 and $+50$ daPa), (2) pure-tone hearing thresholds ≤ 20 dB HL averaged across 500, 1000, 2000 and 4000 Hz and ≤ 25 dB HL for each of these frequencies, (3) word recognition scores in quiet $> 90\%$ (to verify the children's proficiency to repeat simple words under favourable conditions), and (4) no known cognitive or language problems. Fulfilment of these criteria was assessed by means of tympanometry, standard pure-tone audiometry, monosyllabic word recognition measurements in quiet using the DANTALE-I test (Elberling, Ludvigsen, and Lyregaard 1989), and a customised parental questionnaire. The questionnaire included questions related to the child's first language, if the child was monolingual, if the child had shown normal cognitive and language development, and the level of parental education and income. Analysis of these data revealed that all participants were monolingual native Danish speakers with inconspicuous cognitive and language development, and that they came from

families with higher educations as well as middle-to-high incomes. Two Mann-Whitney *U*-tests confirmed no differences in educational status or income of the children's parents between the OM and control group (both $p > 0.6$). Furthermore, according to their parents, none of the participating children had previously been diagnosed with auditory processing disorder.

In addition to the inclusion criteria described above, all children from the OM group were required to have had at least two episodes of middle-ear infection or effusion in at least one ear before age 5.

2.2. Otologic records

For each child in the OM group, the history of the middle-ear diseases was verified based on the medical records from the responsible otologist. The collected records stemmed from several otologists in the Region of Southern Denmark, which is why the information available in the records of the children from the OM group differed. However, all collected records contained information about the number of OM episodes, the age at which these episodes occurred and how long they lasted, and the results of tympanometry and otoacoustic emission measurements (i.e. pass or refer) during and after the disease. Thus, this information served as the basis for the current study. All OM episodes included type-B/C2 tympanograms with normal external ear canal volume as well as "refer" outcomes from otoacoustic emission testing.

All 42 children in the OM group had experienced recurrent acute OM or OM with effusion in both ears, starting with either the left ear, the right ear, or both ears at the same time. Furthermore, all of them had received ventilation tube (VT) treatment at least once. The information contained in the otologic records was systematised by extracting the following information from them: (1) the child's age at the time of the first OM episode ("OM onset age"), (2) the total duration for which a given child had experienced middle-ear disease ("overall OM duration"), and (3) the time interval between the last OM episode and the time of testing ("OM recovery period"). Table 1 provides a summary of these data.

2.3. Speech recognition measurements

To assess the speech recognition abilities of the participants, 50%-correct SRTs were measured. The measurements were controlled via customised MATLAB scripts that simulated a three-dimensional listening environment under headphones. This was achieved by convolving the stimuli with generic anechoic head-related impulse responses measured on a KEMAR mannikin (Gardner and Martin 1995). Recent research has shown that this approach allows for accurate SRT measurements in school-age children (Peng and Litovsky 2021; Zenke and Rosen 2022). The stimuli were presented via an RME Fireface UC soundcard and free-field-equalized Sennheiser HDA200 headphones. The target speech consisted of the test lists from the børneDAT corpus (Koiek et al. 2020). These lists contain 20 sentences each that are

presented in a fixed order. All sentences have a simple, fixed structure, that is, they start with a name (Dagmar, Asta, or Tine) and contain two short, unique keywords. An example is "Dagmar tænkte på en *teske* og en *bjørn* i går" ("Dagmar thought about a *teaspoon* and a *bear* yesterday"). The sentences are uttered by three female speakers (Nielsen, Dau, and Neher 2014). For any given test list, all sentences are uttered by the same speaker, and the initial name (Dagmar, Asta, or Tine) is constant (Koiek et al. 2020). The participants were instructed to pay attention to the sentence starting with a specific name and to repeat the two keywords in that sentence.

The SRT measurements were performed under different acoustic conditions, as explained below. For all measurements, the masker level was fixed at 65 dB SPL and the target speech level was varied according to the adaptive procedure of the Danish Hearing in Noise test (HINT; Nielsen and Dau 2011). From the first to the fourth trial, the target level was decreased by 4 dB if both keywords were repeated correctly; otherwise, a 4-dB increase was applied. From the fifth trial onwards, the step size was reduced to 2 dB. Following the presentation of 20 sentences, the SRT was calculated by averaging the SNRs from the fifth to the hypothetical 21st trial, as done in the Danish HINT. All measurements were conducted in a large sound-attenuating booth. The participants were given a short break after finishing half of the measurements and whenever they felt tired.

2.3.1. Speech recognition in stationary speech-shaped noise (SSN)

SRTs were measured in SSN with an initial target level of 68 dB SPL. The SSN always had the same long-term average speech spectrum as the presented target speech. The target speech was presented from in front (0° azimuth) and SSN from the side (90° azimuth) of the listener. The SRTs were measured either binaurally ("binaural SRTs") or monaurally with the stimuli presented only to the ear opposite the noise, that is, the "better ear" ("monaural SRTs") (Figure 1). For training purposes, the participants performed one binaural SRT measurement in quiet and one binaural SRT measurement in noise. The data from these training measurements were not included in the analyses.

2.3.2. Speech recognition in two-talker running speech (TTS)

SRTs were also measured in TTS. In that case, the initial target level was 72 dB SPL. The target sentences were presented from in front (0° azimuth). As interferers, two stories uttered by two female Danish talkers taken from the DANTALE-I corpus (Elberling, Ludvigsen, and Lyregaard 1989) and the Archimedes project (Hansen and Munch 1991) were used. Consequently, the voice characteristics (fundamental frequency and spectral shape) of the three concurrent talkers were different. SRTs were measured in a co-located and a spatially separated condition. In the co-located condition, the two competing talkers were presented from in front (0° azimuth; "co-located SRTs"), while in the spatially separated condition they were presented from the sides ($\pm 90^\circ$ azimuth; "spatially separated SRTs") (Figure 1). For training purposes, the participants performed one SRT measurement in the spatially separated condition. The data from these training measurements were not included in the analyses.

Table 1. Summary of the three otologic history factors.

Otologic history factor	Mean (SD)	Range
Overall OM duration (years)	0.9 (0.5)	0.1-2.1
OM onset age (years)	1.1 (0.7)	0.3-3.0
OM recovery period (years)	5.4 (3.2)	0.3-10.8

Values in parentheses correspond to standard deviations (SD).

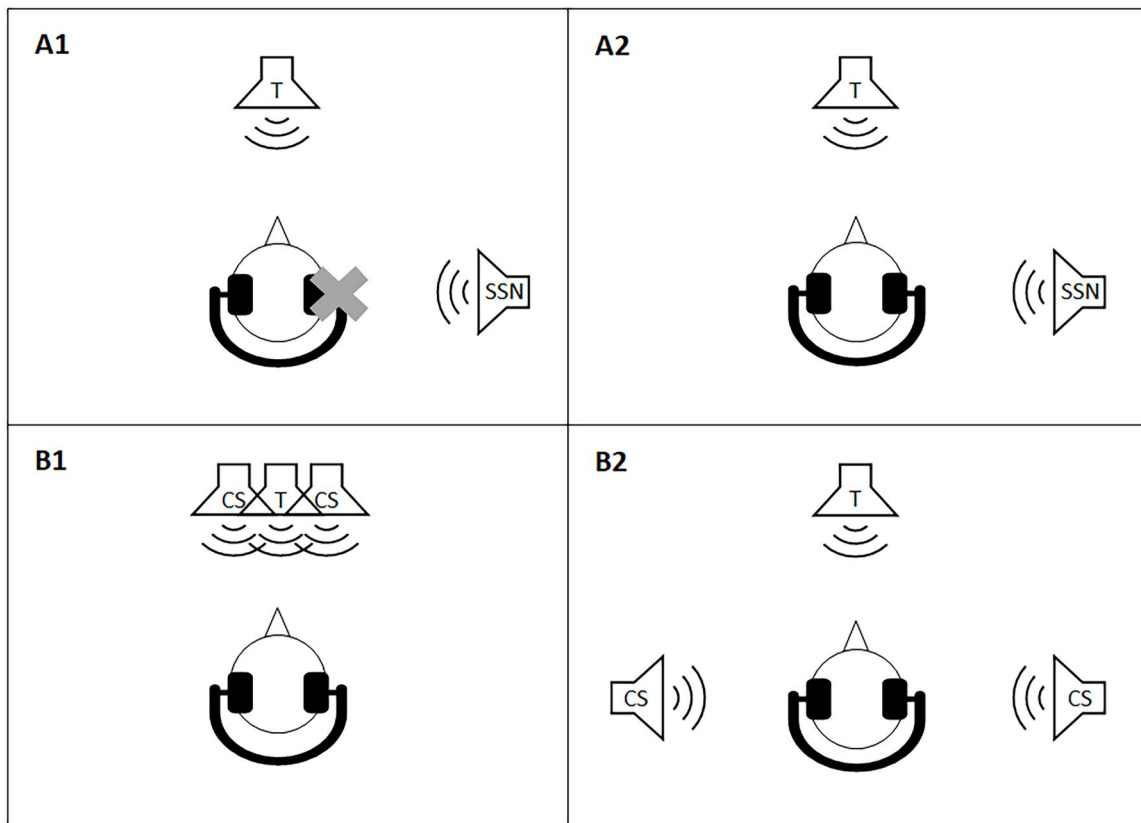


Figure 1. Illustration of the four acoustic conditions. Target speech (T) against SSN with T from in front and SSN from the side, presented either to the “better ear” (A1, monaural condition) or binaurally (A2, binaural condition). T against TTS with T from in front and two competing speech (CS) signals from either in front (B1, co-located condition) or the two sides (B2, spatially separated condition). In all four conditions, stimulus presentation was via headphones. In condition A1 the participants listened monaurally, while in conditions A2, B1, and B2 they listened binaurally.

2.4. Advantage scores

Based on the SRTs for the four acoustic conditions, two types of advantage scores were calculated: (1) *Binaural advantage* scores, which were derived by subtracting the binaural SRTs from the monaural SRTs, and (2) *spatial advantage* scores, which were derived by subtracting the spatially separated SRTs from the co-located SRTs. In the research literature, the terms “binaural intelligibility level difference” (e.g. Neher, 2017) and “spatial release from masking” (e.g. Peng, Pausch, and Fels 2021) respectively are also used for these measures.

Masked speech recognition is facilitated by different types of auditory cues, including monaural head-shadow effects, binaural redundancy, and interaural time and level differences (e.g. Dieudonné and Francart 2019; Peng and Litovsky 2021). As derived here, the binaural advantage scores reflect primarily the ability to benefit from interaural time and level differences in the presence of stationary noise (SSN). Correspondingly, the spatial advantage scores reflect primarily the ability to benefit from interaural time and level differences in the presence of competing speech (TTS). Using these two measures, the binaural contribution to the ability to recognise speech masked by either SSN or TTS was assessed for the two participant groups.

2.5. General procedure

The measurements were completed at two visits lasting for 40–50 min each. First, all participants were screened to verify that they fulfilled the inclusion criteria. Next, the speech recognition measurements were performed. For each participant and acoustic

condition, a set of test and retest measurements was completed at the same visit. If a given retest measurement deviated by more than 3 dB from the corresponding test measurement, another (“repeat”) measurement was conducted. In the data analyses, the median of each set of SRT measurements was used. The order of the four acoustic conditions was randomised across the participants, as were the test lists from the DAT corpus used for the SRT measurements.

2.6. Statistical analyses

The statistical analyses were conducted using Stata version 17 (stataBE 17). In all cases, a significance level of 5% ($\alpha=0.05$) was used.

To begin with, the raw data were inspected with the aim to find any spurious datapoints. First, the standard deviation (SD) of the test, retest, and repeat measurements for each measure (e.g. the monaural SRTs) and participant was calculated. Next, boxplots of all the participants’ SDs were made to identify outliers based on Tukey’s (1977) well-known upper and lower “fences”. In that manner, datapoints more than 1.5 times the interquartile range above the third quartile or below the first quartile of a given dataset were identified. The data from participants with SDs exceeding these fences were excluded from all subsequent analyses. The remaining data were considered genuine observations and were kept. Out of more than 500 measured SRTs, one monaural SRT, two co-located SRTs, and two spatially separated SRTs from four children from the OM group were removed (i.e. <1% of all collected SRTs). As a result, complete datasets were available from 38 children from the OM group.

Following the data cleaning, the SRTs and advantage scores were analysed in terms of their repeatability. To that end, the within-subject SD was calculated for each measure. The correlation between the test and retest results was also examined using Spearman's ρ correlation coefficient. Participants with more than two measurements (test, retest, repeat) were excluded from this analysis. The within-subject SD was found to be <1.3 dB for all measures. Moreover, except for the binaural advantage scores ($\rho = 0.25$, $p = 0.052$), ρ ranged from 0.65 to 0.95 (all $p < 0.001$), indicating good repeatability overall.

To examine the distributions of the collected datasets, Shapiro-Wilk's test, normal Q-Q plots, and boxplots were used. To verify the equality of variances, Levene's test was used. Importantly, outliers (see above) were not excluded at this stage because they were expected in the data of the OM group and were therefore considered genuine observations.

To examine the effects of participant group and acoustic condition/advantage type, two mixed-effects linear regression models with repeated measures were performed, that is, one based on the SRTs (monaural, binaural, co-located, spatially separated) and one based on the advantage scores (binaural, spatial). Both models included a random intercept per participant as well as an interaction term (participant group \times acoustic condition/advantage type). In addition, they included age as a covariate to control for any changes in developmental abilities that typically accompany higher age (e.g. improvements in masked speech recognition). The quality of the models was examined by checking the residuals and random effects. In that manner, the assumptions of homoscedasticity and normality were confirmed.

To examine the influence of the otologic history factors (overall OM duration, OM onset age, OM recovery period), six linear regression analyses with age and the three otologic history factors as predictors were performed. As dependent variables, the four types of SRTs and two types of advantage scores were used.

3. Results

3.1. SRT measurements

Figure 2 shows mean SRTs together with 95% confidence intervals (CIs) for the four acoustic conditions and two participant groups. Individual datapoints are also included. The mixed-

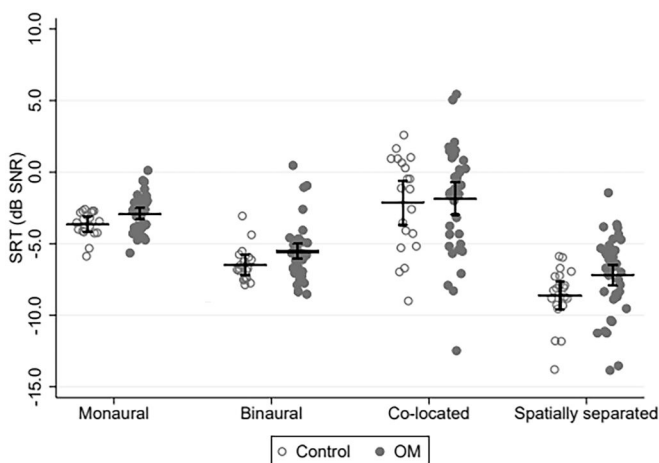


Figure 2. Mean speech recognition thresholds (SRTs) with 95% CIs and individual datapoints for the four acoustic conditions. In the monaural and binaural conditions, SSN was used. In the co-located and spatially separated conditions, TTS was used. Unfilled circles show data from the controls. Filled circles show data from the children with a history of otitis media (OM).

effects linear regression model revealed effects of acoustic condition ($p = 0.007$) and participant group (Δ -value = 0.84 dB; $p = 0.009$; 95% CI: [0.22, 1.47] dB). The interaction between acoustic condition and participant group exceeded the 5% significance level slightly ($p = 0.053$).

The linear regression analyses revealed a significant influence of age on the monaural, binaural, and spatially separated SRTs (all $p < 0.043$; Table 2) but not on the co-located SRTs ($p = 0.14$). Overall OM duration, OM onset age, and OM recovery period did not show up as significant predictors in any of the models tested (all $p > 0.2$).

3.2. Advantage scores

Figure 3 shows mean binaural and spatial advantage scores together with 95% CIs for the control and OM groups. Individual scores are also included. The mixed-effects linear regression model showed an effect of advantage type ($p < 0.001$) but not participant group ($p = 0.37$). The linear regression analyses revealed effects of neither age nor the three otologic history factors (all $p > 0.3$).

3.3. Follow-up group comparisons

To shed more light on the group difference and the trend for an interaction between acoustic condition and participant group observed above, exploratory group comparisons were performed on the SRTs from the four acoustic conditions. These revealed that the monaural, binaural, and spatially separated SRTs were elevated by, respectively, 0.75, 0.98, and 1.44 dB for the OM group relative to the controls (all $p < 0.023$). Table 3 provides a summary of these results.

4. Discussion

The current study investigated the long-term effects of early-childhood OM on speech recognition in the presence of SSN or TTS with or without interaural differences among the competing signals. Groups of school-age children with a history of OM or without any ear diseases were tested. The influence of three otologic history factors (overall OM duration, OM onset age, and OM recovery period) was also investigated. Overall, the OM group showed elevated SRTs compared to the controls (Δ -value = 0.84 dB, $p = 0.009$), which appeared driven by the spatially separated, binaural, and monaural conditions. No group effects were apparent in the advantage scores. In general, there were large inter-individual differences, particularly so for the OM group, which were unrelated to the otologic history factors. Below, these findings are discussed.

Table 2. Results of linear regression analyses performed with the four types of SRTs as dependent variables and age, overall OM duration, OM onset age, and OM recovery period as independent variables.

Condition	R^2	Adj. R^2	Sig. predictor	β -value (95% CI)	p -value
Monaural	0.12	0.02	Age	-0.73 (-1.44, -0.03)	0.042
Binaural	0.25	0.16	Age	-0.39 (-0.74, -0.04)	0.031
Spatially sep.	0.42	0.35	Age	-0.74 (-1.17, -0.31)	0.001

CI: Confidence interval.

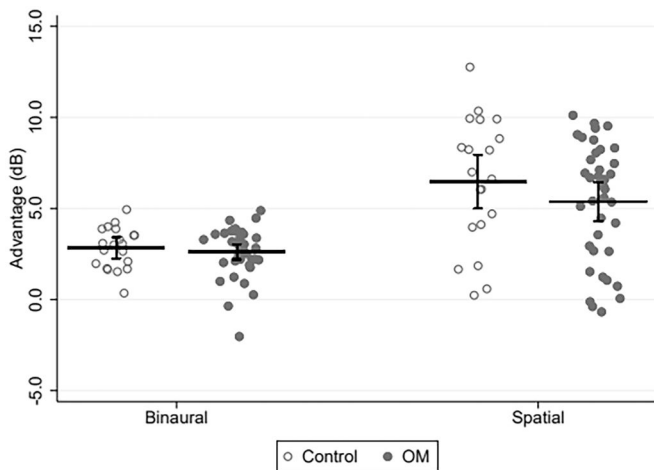


Figure 3. Mean binaural and spatial advantage scores with 95% CIs as well as individual datapoints for the controls (unfilled circles) and the children with a history of otitis media (OM; filled circles). For the binaural advantage measurements, SSN was used. For the spatial advantage measurements, TTS was used.

Table 3. Results from follow-up group comparisons based on mixed-effects linear regression analyses with age as covariate and a random intercept per participant.

SRT (dB SNR)	Mean (SD)		Mean difference (95% CI)	<i>p</i> -value
	Controls	OM group		
Monaural	-3.62 (0.89)	-2.92 (1.32)	0.75 (0.11, 1.40)	0.023
Binaural	-6.46 (1.17)	-5.53 (1.98)	0.98 (0.16, 1.79)	0.020
Co-located	-2.14 (3.31)	-1.85 (3.70)	0.34 (-1.54, 2.21)	0.73
Spatially sep.	-8.61 (1.99)	-7.21 (2.76)	1.44 (0.28, 2.60)	0.015

SD: Standard deviation. CI: Confidence interval.

4.1. Influence of acoustic condition

The current study found an effect of acoustic condition on the SRTs. For the measurements in SSN, the mean SRT improved significantly when the signals were presented binaurally (and interaural differences were available) compared to when they were presented monaurally (and interaural differences were unavailable). Correspondingly, for the measurements in TTS, the mean SRT was significantly better when the target and masker signals were spatially separated relative to when they were co-located with the target speech. In other words, the groups tested here had sufficiently intact binaural hearing abilities to benefit from the availability of interaural differences among the competing sound signals.

Overall, the SRTs were highest in the co-located condition and lowest in the spatially separated condition (Figure 2), resulting in greater spatial advantage than binaural advantage (overall means: 5.7 dB vs. 2.7 dB; Figure 3). Furthermore, inter-individual differences were larger in the presence of TTS than in the presence of SSN (Figures 2–3). These differences, which can be attributed to informational masking effects arising from the use of the TTS masker, fit well with what is known about masked speech recognition in children and adults (e.g. Bronkhorst 2000; Yuen and Yuan 2014).

4.2. Influence of OM status

The current study found an adverse effect of early-childhood OM on masked speech recognition, and a trend for an interaction between participant group and acoustic condition ($p=0.053$). Group effects were clearest in the spatially separated

condition (Δ -value = 1.44 dB) followed by the binaural condition (Δ -value = 0.98 dB), that is, when interaural differences were available. These results are comparable to the group difference found by Graydon et al. (2017) under spatially separated conditions using the LiSN-S (see Introduction). While this would seem to lend support to the idea that early-childhood OM is detrimental for binaural hearing abilities, the advantage scores did not support such a conclusion (see below).

The overall group difference found here was small (Δ -value = 0.84 dB). Indeed, most children in the OM group obtained results in the normal range (Figure 2). Nevertheless, some children in the OM group showed noticeably poorer performance, with a few obtaining negative advantage scores (Figure 3). It is possible that differences in treatment history could explain these findings. According to Lous et al. (1999), children with OM who receive VT treatment spend on average 32% less time with effusion and the accompanying hearing loss compared to untreated children. Perhaps those children with OM whose results were in the normal range received VT treatment sooner as compared with those children with OM who showed abnormal results. The available otologic records (Sect. 2.2) did not contain information about the VT treatment of all children from the OM group, and so this possibility would have to be investigated in follow-up research.

Compared to SRTs, advantage scores arguably allow for a more precise examination of the role of binaural hearing since factors related to, for example, monaural processing, language, and cognition are effectively factored out (Dillon and Cameron 2021). The current study did not find any effects of OM status on binaural or spatial advantage. One explanation for this could be the greater variability that is typical of advantage scores as compared with SRTs (Neher, Fogh, and Koiek 2022). In general, difference scores are affected by more random measurement error, as the random error in the two underlying conditions adds up in the difference measure. Nevertheless, a previous study (Koiek et al. 2022) found an effect of OM status on the binaural masking level difference – a differential measure based on diotic and dichotic tone-in-noise detection thresholds (e.g. Neher, 2017). This could suggest that the binaural masking level difference is more sensitive to OM-related hearing deficits than (speech-based) advantage scores.

Another explanation could be related to the voice characteristics of the target and competing speech signals used in the current study. In the study by Cameron and Dillon (2007), normal-hearing children showed a mean spatial advantage of 12.1 dB in the “same voice” condition and of 9.8 dB in the “different voices” condition. In the current study, which made use of three different speech materials, spatial advantage was on average 6.5 dB for the controls. The smaller spatial advantage observed here could be due to the use of different talkers (with different fundamental frequencies and spectral shapes) for the target speech and speech interferers. When such differences are more pronounced, spatial cues play a smaller role for task performance (e.g. Bronkhorst 2000). This, in turn, would make it less likely to find a deficit in the ability to exploit interaural differences because of early-childhood OM.

The lack of a group difference in terms of spatial advantage contrasts with the findings of Tomlin and Rance (2014) and Graydon et al. (2017) who observed reduced spatial advantage in school-age children with early-childhood OM (see Introduction). A possible explanation for this could be differences between the study samples. As mentioned above, all children in the OM group tested here had received VT treatment at least once. VT

treatment can be expected to alleviate the effects of auditory deprivation due to CHL. It is thus possible that the effects of early-childhood OM observed here are less pronounced compared to studies with children with an OM history from countries where VT treatment is less common (Pedersen et al. 2016). This explanation would be in line with results of Borges et al. (2020) who found that the long-term effects of early-childhood OM on auditory processing depend on the children's country of residence. In that study, temporal processing in Brazilian school-age children with a documented OM history before age 6 was significantly poorer compared to Australian peers with a similar otologic history. Given that much OM-related research has been conducted in developed countries where medical intervention is the norm, future efforts should be directed at geographical regions with less developed healthcare systems where the effects of early-childhood OM on masked speech recognition may be more severe.

It is also possible that the time interval between the last OM episode and the time of testing was longer in the current study (mean: 5.4 years \pm 3.2 years) compared with previous studies. If so, the children tested here would have had more time to recover from their hearing deficits, thereby weakening potential group effects. Such an explanation would be consistent with the results of Hall, Grose, and Pillsbury (1995) who suggested that long-term binaural hearing impairments caused by early-childhood OM tend to disappear three years after VT treatment.

4.3. Influence of otologic history factors

The current study did not find any effects of overall OM duration, OM onset age, or OM recovery period on long-term auditory outcome. Some (Tomlin and Rance 2014; Zumach et al. 2009) but not all (Keogh et al. 2005) literature findings are at odds with this result. Gravel and Wallace (1992, 1995) found that the adverse effects of OM on language development, auditory processing, and academic achievement were associated with the *degree* of hearing loss rather than the presence of OM. Likewise, Zumach et al. (2011), who tested 7-year-olds with documented OM from birth to 24 months of age, proposed that it is not OM per se but rather the severity of CHL that is related to poorer speech perception abilities in such children. Thus, differences in CHL severity can perhaps explain the discrepant findings across studies.

The ineffectiveness of the three otologic history factors to explain the large inter-individual variability in the OM group could have been due to the retrospective study design. With this type of design, there is a risk of missing important or obtaining inaccurate participant information, for example the exact dates of OM on- and offset (Hartley and Moore 2005; Zumach et al. 2009). What is more, there is a risk of recruitment bias. The use of otologic records only captures patients seeking medical care for obvious OM. Consequently, many cases of OM go undetected, which is a problem for OM studies in general. Overall, a prospective study that controls for these risks would be suited for investigating the influence of the otologic history factors on auditory abilities.

4.4. Limitations and future directions

For the controls, the otologic history was assessed based on parental reports only. Children in this group could have had undiagnosed middle-ear diseases during early childhood, leading to an underestimation of group differences. Also, as mentioned previously, the collected otologic records varied in scope and

lacked information about the OM group's hearing thresholds during episodes of the disease. Despite this, the current study found a significant group difference in masked speech recognition, which suggests that the group definition was robust. On average, this difference was <1 dB, implying little, if any, real-world impact. There were, however, children whose SRTs were more than 5 dB worse than the means of the control group (see Figure 2, binaural and spatially separated conditions), suggesting large effects at the individual level. Future work should follow up on the clinical significance of these results. Ideally, such efforts should also lead to clinical recommendations related to the audiologic and medical management of the disease.

The large inter-individual variability in the OM group was unlikely related to parental education level, socioeconomic status, language, or cognitive skills. To recapitulate, all participants came from families with higher educations as well as middle-to-high incomes (Sect. 2.1). Also, the target sentences used for the speech recognition measurements had a simple, fixed structure with two short keywords each (Sect. 2.3). Instead, the large inter-individual variability could have been related to differences in peripheral function such as high-frequency hearing thresholds (Petley et al. 2021) or the ability to use grouping and segregation cues, especially in the context of competing-speech stimuli (Bronkhorst 2000). The current study did not assess these functions and abilities directly. Ideally, future work should do this, so these types of influences on the speech recognition abilities of children with an OM history can be accounted for as well.

Lastly, a longitudinal study design would open avenues for examining the time course of OM-induced hearing deficits. This could reveal if the effects of OM on speech recognition persist for a long time or if they are more acute in nature. Moreover, it could enable a more fine-grained analysis of the predictive power of the otologic history factors. For example, it is possible that the progression of the disease matters most for a child's hearing abilities in the first years after OM resolution and less so afterwards.

Acknowledgements

The authors thank Signe Hjorth Fogh and Nadia Føns Christensen for help with the data collection as well as all the participants and parents for their efforts. Author HD acknowledges support of the NIHR Manchester Biomedical Research Centre.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research was supported by a PhD stipend from the William Demant Foundation (case no. 17-4167). The funding source had no role in the study design, the collection, analysis, or interpretation of the data, the writing of the manuscript, or in the decision to submit the article for publication.

ORCID

Shno Koiek  <http://orcid.org/0000-0001-5779-3599>
 Christian Brandt  <http://orcid.org/0000-0003-0672-9487>
 Sören Möller  <http://orcid.org/0000-0003-0858-4269>
 Harvey Dillon  <http://orcid.org/0000-0002-7036-7890>
 Tobias Neher  <http://orcid.org/0000-0002-1107-9274>

Data availability statement

The data and code used for the statistical analyses can be downloaded from <https://doi.org/10.5281/zenodo.8409629>.

References

- Borges, L. R., M. D. Sanfins, C. Donadon, D. Tomlin, and M. F. Colella-Santos. 2020. "Long-Term Effect of Middle Ear Disease on Temporal Processing and P300 in Two Different Populations of Children." *PloS One* 15 (5):e0232839. <https://doi.org/10.1371/journal.pone.0232839>.
- Bronkhorst, A. W. 2000. "The Cocktail Party Phenomenon: A Review of Research on Speech Intelligibility in Multiple-Talker Conditions." *Acta Acustica United with Acustica* 86 (1):117–128.
- Cameron, S., and H. Dillon. 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>.
- Dawes, P., and D. Bishop. 2009. "Auditory Processing Disorder in Relation to Developmental Disorders of Language, Communication, and Attention: A Review and Critique." *International Journal of Language & Communication Disorders* 44 (4):440–465. <https://doi.org/10.1080/13682820902929073>.
- Dieudonné, B., and T. Francart. 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.00000000000000060>.
- Dillon, H., and S. Cameron. 2021. "Separating the Causes of Listening Difficulties in Children." *Ear and Hearing* 42 (5):1097–1108. <https://doi.org/10.1097/AUD.0000000000001069>.
- Elberling, C., C. Ludvigsen, and P. E. Lyregaard. 1989. "DANTALE: A New Danish Speech Material." *Scandinavian Audiology* 18 (3):169–175. <https://doi.org/10.3109/01050398909070742>.
- Gardner, W. G., and K. D. Martin. 1995. "HRTF Measurements of a KEMAR." *The Journal of the Acoustical Society of America* 97 (6):3907–3908. <https://doi.org/10.1121/1.412407>.
- Gravel, J. S., and I. F. Wallace. 1992. "Listening and Language at 4 Years of Age." *Journal of Speech and Hearing Research* 35 (3):588–595. <https://doi.org/10.1044/jshr.3503.588>.
- Gravel, J. S., and I. F. Wallace. 1995. "Early Otitis Media, Auditory Abilities, and Educational Risk." *American Journal of Speech-Language Pathology* 4 (3):89–94. <https://doi.org/10.1044/1058-0360.0403.89>.
- Graydon, K., G. Rance, R. Dowell, and B. Van Dun. 2017. "Consequences of Early Conductive Hearing Loss on Long-Term Binaural Processing." *Ear and Hearing* 38 (5):621–627. <https://doi.org/10.1097/aud.0000000000000431>.
- Haggard, M., and E. Hughes. 1991. *Screening Children's Hearing: A Review of the Literature and the Implications of Otitis Media*. HM Stationery Office.
- Hall, J. W., J. H. Grose, and H. C. Pillsbury. 1995. "Long-Term Effects of Chronic Otitis Media on Binaural Hearing in Children." *Archives of Otolaryngology-Head & Neck Surgery* 121 (8):847–852. <https://doi.org/10.1001/archotol.1995.01890080017003>.
- Hansen, V., and G. Munch. 1991. "Making Recordings for Simulation Tests in the Archimedes Project." *Journal of the Audio Engineering Society* 39 (10):768–774.
- Hartley, D. E. H., and D. R. Moore. 2005. "Effects of Otitis Media with Effusion on Auditory Temporal Resolution." *International Journal of Pediatric Otorhinolaryngology* 69 (6):757–769. <https://doi.org/10.1016/j.ijporl.2005.01.009>.
- Keogh, T., J. Kei, C. Driscoll, L. Cahill, A. Hoffmann, E. Wilce, P. Kondamuri, and J. Marinac. 2005. "Measuring the Ability of School Children with a History of Otitis Media to Understand Everyday Speech." *Journal of the American Academy of Audiology* 16 (5):301–311. <https://doi.org/10.3766/jaaa.16.5.5>.
- Koiek, S., C. Brandt, J. H. Schmidt, and T. Neher. 2022. "Monaural and Binaural Phase Sensitivity in School-Age Children with Early-Childhood Otitis Media." *International Journal of Audiology* 61 (12):1054–1061. <https://doi.org/10.1080/14992027.2021.2009132>.
- Koiek, S., J. B. Nielsen, L. Kjørnbæk, M. B. Gormsen, and T. Neher. 2020. "A Danish Sentence Corpus for Assessing Speech Recognition in Noise in School-Age Children." *Trends in Hearing* 24:2331216520942392. <https://doi.org/10.1177/2331216520942392>.
- Lous, J., M. Burton, J. Felding, T. Ovesen, M. Wake, and I. Williamson. 1999. "Grommets (ventilation tubes) for Hearing Loss Associated with Otitis Media with Effusion in Children." *The Cochrane Database of Systematic Reviews* 2010 (10):CD001801. <https://doi.org/10.1002/14651858.CD001801.pub3>.
- Moore, D. R., D. E. Hartley, and S. C. Hogan. 2003. "Effects of Otitis Media with Effusion (OME) on Central Auditory Function." *International Journal of Pediatric Otorhinolaryngology* 67 (Suppl. 1):S63–S67. <https://doi.org/10.1016/j.ijporl.2003.08.015>.
- Neher, T. 2017. "Characterizing the Binaural Contribution to Speech-In-Noise Reception in Elderly Hearing-Impaired Listeners." *The Journal of the Acoustical Society of America* 141 (2):EL159–EL163. <https://doi.org/10.1121/1.4976327>.
- Neher, T., S. H. Fogh, and S. Koiek. 2022. "Masked Speech Recognition by Normal-Hearing 6-13-Year-Olds in Conditions With and Without Interaural Difference Cues." *Trends in Hearing* 26:23312165221137117. <https://doi.org/10.1177/23312165221137117>.
- Nielsen, J. B., and T. Dau. 2011. "The Danish Hearing in Noise Test." *International Journal of Audiology* 50 (3):202–208. <https://doi.org/10.3109/14992027.2010.524254>.
- Nielsen, J. B., T. Dau, and T. Neher. 2014. "A Danish Open-Set Speech Corpus for Competing-Speech Studies." *The Journal of the Acoustical Society of America* 135 (1):407–420. <https://doi.org/10.1121/1.4835935>.
- Pedersen, T. M., A.-R. C. Mora-Jensen, J. Waage, H. Bisgaard, and J. Stokholm. 2016. "Incidence and Determinants of Ventilation Tubes in Denmark." *PloS One* 11 (11):e0165657. <https://doi.org/10.1371/journal.pone.0165657>.
- Peng, Z. E., and R. Y. Litovsky. 2021. "The Role of Interaural Differences, Head Shadow, and Binaural Redundancy in Binaural Intelligibility Benefits Among School-Aged Children." *Trends in Hearing* 25:23312165211045313. <https://doi.org/10.1177/23312165211045313>.
- Peng, Z. E., F. Pausch, and J. Fels. 2021. "Spatial Release From Masking in Reverberation for School-Age Children." *The Journal of the Acoustical Society of America* 150 (5):3263–3274. <https://doi.org/10.1121/10.0006752>.
- Petley, L., L. L. Hunter, L. Motlagh Zadeh, H. J. Stewart, N. T. Sloat, A. Perdew, L. Lin, and D. R. Moore. 2021. "Listening Difficulties in Children with Normal Audiograms: Relation to Hearing and Cognition." *Ear and Hearing* 42 (6):1640–1655. <https://doi.org/10.1097/AUD.0000000000001076>.
- Tomlin, D., and G. Rance. 2014. "Long-Term Hearing Deficits after Childhood Middle Ear Disease." *Ear and Hearing* 35 (6):e233–e242. <https://doi.org/10.1097/aud.0000000000000065>.
- Tukey, J. W. 1977. *Exploratory Data Analysis*. Reading, MA: Addison-Wesley.
- Whitton, J. P., and D. B. Polley. 2011. "Evaluating the Perceptual and Pathophysiological Consequences of Auditory Deprivation in Early Postnatal Life: A Comparison of Basic and Clinical Studies." *Journal of the Association for Research in Otolaryngology: JARO* 12 (5):535–547. <https://doi.org/10.1007/s10162-011-0271-6>.
- Yuen, K. C. P., and M. Yuan. 2014. "Development of Spatial Release From Masking in Mandarin-Speaking Children with Normal Hearing." *Journal of Speech, Language, and Hearing Research: JSLHR* 57 (5):2005–2023. https://doi.org/10.1044/2014_JSLHR-H-13-0060.
- Zenke, K., and S. Rosen. 2022. "Spatial Release of Masking in Children and Adults in Non-Individualized Virtual Environments." *The Journal of the Acoustical Society of America* 152 (6):3384–3395. <https://doi.org/10.1121/10.0016360>.
- Zumach, A., M. N. Chenault, L. J. C. Anteunis, and E. Gerrits. 2011. "Speech Perception After Early-Life Otitis Media with Fluctuating Hearing Loss." *Audiology & Neuro-Otology* 16 (5):304–314. <https://doi.org/10.1159/000322501>.
- Zumach, A., E. Gerrits, M. N. Chenault, and L. J. C. Anteunis. 2009. "Otitis Media and Speech-In-Noise Recognition in School-Aged Children." *Audiology & Neuro-Otology* 14 (2):121–129. <https://doi.org/10.1159/000162664>.