The Effect of Otitis Media With Effusion on Perceptual Masking

Joseph W. Hall, PhD; John H. Grose, PhD; Emily Buss, PhD; Madhu B. Dev, BA; Amelia F. Drake, MD; Harold C. Pillsbury, MD

Objective: To determine the effect of otitis media with effusion (OME) on perceptual masking (a phenomenon in which spondee threshold for a 2-talker masker is poorer than for a speech-shaped noise masker).

Design: Longitudinal testing over a 1-year period following insertion of tympanostomy tubes, using clinical and normal-hearing control groups.

Subjects: Forty-seven children having a history of OME were tested. Possible testing intervals were just before the placement of tympanostomy tubes, and up to 3 separate occasions after the placement of the tubes. An age-matched control group of 19 children was tested.

Methods: A perceptual masking paradigm was used to measure the ability of the listener to recognize a spondee in either a speech-shaped noise or a 2-talker masker background. The masker was either continuous or gated on and off with the target spondee.

Results: In gated masking conditions, children with a history of normal hearing showed only slight perceptual masking, but the children with a history of OME showed relatively great perceptual masking before surgery and up to 6 months following surgery. In continuous masking conditions, both groups of children showed relatively great perceptual masking and did not differ significantly from each other in this respect either before or after surgery. However, before surgery, the OME group showed higher thresholds in both the 2-talker and speech-shaped noise maskers.

Conclusions: In agreement with previous psychoacoustical findings, the relatively great perceptual masking in gated conditions shown by children with OME history may reflect a general deficit in complex auditory processing.


This study investigated masked speech recognition in children having a history of otitis media with effusion (OME). Although previous studies investigating the effect of OME on the recognition of speech have been somewhat sparse, the available evidence suggests that OME history may be associated with a reduced ability to recognize speech in the presence of a masker. For example, studies of children with a history of OME have shown poor recognition for words in sentences masked by a competing talker, and slightly poorer identification of monosyllabic words masked by speech-shaped noise. Other investigations have also suggested that a history of OME may be associated with poor perception of particular features of speech. A specific approach of the present study was to compare masked speech recognition thresholds (SRTs) between 2 masker types: a speech-shaped noise and a masker composed of 2 competing talkers. It was hypothesized that the effect of OME might be greater in the 2-talker masker owing to factors associated with processing complexity. This hypothesis was based upon previous work suggesting that processing complexity is greater for a 2-talker masker than for a speech-shaped masker (see below), and psychoacoustical results suggesting that effects of OME may be more likely to occur in conditions requiring relatively complex auditory processing. Previous psychoacoustical studies have indicated no effect of OME history on the simple task of diotic detection of a tone in a band of noise. However, a deleterious effect of OME history has been found for masking tasks based upon the coding of binaural difference cues, and the task of comodulation masking release, which is presum-
Continuous and gated conditions. Interestingly, we found that both adults and children had a significant release from perceptual masking when the masker was gated. By including the gating condition in the present study, we can assess the effect of OME on perceptual masking, and on the release from perceptual masking associated with the gated condition.

### METHODS

The control group consisted of 19 children, ranging in age from 5.4 to 10.2 years (mean, 7.5 years). The data from these listeners were reported in the study of Hall et al evaluating developmental effects of perceptual masking. Pure-tone air conduction audiometry indicated that thresholds in quiet were equal to or better than 20 dB hearing level (HL) for octave frequencies between 250 Hz and 8000 Hz. No listeners had a known history of significant ear disease. The experimental group was composed of 47 children ranging in age from 5.1 to 10.9 years (mean, 7.1 years) who had a history of OME.

In this study, only subjects having documented hearing loss of 25 dB HL or worse at 1 or more frequencies between 250 and 2000 Hz and a type B tympanogram in both ears were included in the experimental group. In addition, the presence of OME was supported by otoscopy performed by an otolaryngologist. Audiometric pure-tone thresholds were obtained for each ear, using the descending Hughson-Westlake method.

### SUBJECTS

Most of the children with OME history were tested both just before the placement of tympanostomy tubes, and on as many as 3 separate occasions after the placement of the tubes: 1 month, 4 to 6 months, or 12 to 13 months after tube placement. Many of the children were able to participate in only a subset of the testing intervals (see Table 1 for subjects associated with the various experimental conditions). Because the children in the OME group were tested in conjunction with their clinical visits, their testing was subject to time constraints, and these children were therefore assigned either to the gated masking conditions or to the continuous masking conditions. The children with a history of normal hearing were tested on both gated and continuous conditions (tested on separate days and randomly assigned as to whether the gated or continuous conditions would be tested on the first or second day). The children with a history of normal hearing were not tested on further occasions because a pilot study indicated that the masked

**Table 1. Mean Pure-Tone Average Thresholds for the Children Tested in the Various Condition Groups and Testing Intervals**

<table>
<thead>
<tr>
<th>Testing Interval</th>
<th>Condition</th>
<th>Left Ear</th>
<th>Right Ear</th>
<th>Mean Perceptual Masking, dB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>Gated (n = 19)</td>
<td>6.1</td>
<td>6.3</td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td>Continuous (n = 19)</td>
<td>6.1</td>
<td>6.1</td>
<td>6.6</td>
</tr>
<tr>
<td>OME presurgery</td>
<td>Gated (n = 12)</td>
<td>24.0</td>
<td>22.3</td>
<td>6.3</td>
</tr>
<tr>
<td></td>
<td>Continuous (n = 21)</td>
<td>25.6</td>
<td>22.4</td>
<td>6.2</td>
</tr>
<tr>
<td>OME 1 mo</td>
<td>Gated (n = 17)</td>
<td>9.2</td>
<td>9.6</td>
<td>6.5</td>
</tr>
<tr>
<td></td>
<td>Continuous (n = 20)</td>
<td>7.1</td>
<td>7.1</td>
<td>6.1</td>
</tr>
<tr>
<td>OME 4-6 mo</td>
<td>Gated (n = 11)</td>
<td>9.5</td>
<td>8.6</td>
<td>5.0</td>
</tr>
<tr>
<td></td>
<td>Continuous (n = 13)</td>
<td>7.4</td>
<td>7.1</td>
<td>5.4</td>
</tr>
<tr>
<td>OME 12-13 mo</td>
<td>Gated (n = 9)</td>
<td>8.3</td>
<td>8.7</td>
<td>2.4</td>
</tr>
<tr>
<td></td>
<td>Continuous (n = 9)</td>
<td>9.4</td>
<td>7.8</td>
<td>6.1</td>
</tr>
</tbody>
</table>

Abbreviations: dB HL, decibels hearing level; OME, otitis media with effusion.

*Frequencies of 500, 1000, and 2000 Hz were used to compute the pure-tone averages.
SRTs, estimated over 2 separate testing days, did not change significantly with practice.

STIMULI

The target speech was taken from a set of 25 digitally recorded spondees (male speaker). The spondees were airplane, arm chair, baseball, bathtub, birthday, bluebird, cowboy, cupcake, doormat, flashlight, football, hotdog, ice cream, mailman, mousetrap, mushroom, playground, popcorn, sailboat, seasaw, shoe lace, sidewalk, snowman, toothbrush, and toothpaste. The listeners were given a "familiarization trial" in which each spondee was presented in quiet while the associated illustration was shown on a monitor.

The speech masker was composed of 2 male talkers, producing continuous, meaningful speech streams. One message was biographical information about Davy Crockett, derived from the Synthetic Sentence Identification test. The other message was composed of 16 sentences from list 1 of the BKB sentences, supplied by a CD from the Cochlear Corporation (Englewood, Colo). The sentences were digitized and played continuously by repeating the sequence of 16 sentences in a "seamless" sequence. The 2 messages were equalized for level, mixed, and then recorded on an audio CD. The long-term average spectrum of the 2-talker masker was used to determine the spectral shape for the speech-shaped noise masker. This spectral shape was then sampled at intervals of 100 Hz from 0 to 6000 Hz, and the resulting values were used to program a digital filter (Tucker-Davis PD1; Tucker-Davis Technologies, Alachua, Fla) using a sampling rate of 50 kHz. A broadband noise source was used to input the PD1. Further details regarding the masker can be found in Hall et al.

The spondees and the speech-shaped noise masker were low-pass filtered at 8 kHz (Kemo VBF8; Kemo Inc, Jacksonville, Fla). Stimuli were passed through switches (Tucker-Davis SW2), with the masker switches left open for the continuous masker conditions.

The masker was played throughout the threshold run in the continuous masker condition. In the gated conditions, the maskers were played exactly as for the continuous condition but were gated on only during spondee presentation. The gated maskers were 1647 milliseconds in duration and were gated on 5 milliseconds before the spondees were gated on. The spondees were 760 to 1150 milliseconds in duration. The masker was therefore gated off 492 to 882 milliseconds after the end of the spondee. A cosine squared rise/fall of 5 milliseconds was used for all gating. Spondees were attenuated adaptively (see below) using a programmable attenuator (Tucker-Davis PA4). The spondee words and the maskers were mixed (Tucker-Davis SM3), sent to a headphone buffer (Tucker-Davis HB6), and delivered to the listeners through headphones (Sony MDR V6). The masker (either 2-talker or speech-shaped noise) was presented at an overall level of 70 dB sound pressure level. The signal and masker were presented diotically.

PROCEDURE

Listeners sat in front of a video display in a double-walled sound booth. On each trial, the listener was presented with a visual display consisting of 4 pictures drawn without replacement from the set of 25. One of the pictures corresponded to the target spondee, and the other 3 were randomly drawn from the list of the remaining possible spondees. Each picture was assigned randomly to 1 of the 4 quadrants of the video display. The pictures were displayed approximately 20 milliseconds before the spondee was presented. The target spondee was chosen randomly on each signal presentation within a threshold run, and the order of spondee presentation was independent across threshold runs. Children pointed to 1 of the 4 spondee illustrations and the accompanying experimenter entered the choice via the keyboard. Visual feedback was provided after each response by highlighting the appropriate picture on the visual display. Data were collected using a 4-alternative, forced-choice, adaptive strategy incorporating a 3-down 1-up stepping rule that estimated the 79.4% correct point on the psychometric function.

Following 3 correct responses in succession, the level of the spondee was reduced; following a single incorrect response, the level of the spondee was increased. An initial step-size of 8 dB was reduced to 4 dB after the first reversal, and further reduced to 2 dB after the second reversal. A threshold run was stopped after 6 reversals, and the average of the final 4 reversals was taken as the SRT for the run. At least 2 SRTs were obtained for each condition, and a third SRT was obtained if the first 2 differed from each other by more than 3 dB. The order of stimulus conditions was random for each listener, but all SRTs for a particular condition were completed before moving on to a new condition. Practice was not provided on the SRT estimation procedure. The above methodological details for stimulus and procedure are identical to those reported in Hall et al.

DATA ANALYSIS

Group and condition differences were analyzed using analysis of covariance (ANCOVA), with age as the covariate, and criterion of significance of $P<.05$.

RESULTS

Individual data for all conditions of the experiment, plotted as a function of age, are summarized in Figure 1 and Figure 2. Each figure contains separate panels showing data obtained for the 2 types of maskers, and the derived measures of perceptual masking (SRT in the speech-shaped noise masker minus SRT in the 2-talker masker). Figure 1 shows the data for continuous masking, and Figure 2 for gated masking. In each figure, the circles depict the data of the children with no known history of OME, and the triangles depict the data of the children with OME. The triangles are below the normal mean for the gated 2-talker masker condition, but that are similar for the postsurgery continuous noise conditions. The results of the study were relatively straightforward and can be summarized with the statement that the performance of the 2 groups of children was generally similar for the postsurgery continuous noise conditions and the gated speech-shaped noise condition, but that the OME-history children often showed performance above the normal mean for the gated 2-talker masker condition. Another trend was a general improvement in masked SRTs as a function of increasing listener age (Figures 1 and 2). Because of this trend, ANCOVA, using age as the covariate, was performed to investigate the significance of masked SRT effects related to experimental group and masking condition. Age was found to have a significant effect in all of these analyses on masked SRTs (Table 2).
CONTINUOUS MASKER

Before surgery, the ANCOVA analysis (Table 2) indicated a significant effect of group (the OME group had significantly poorer SRTs), and a significant effect of masker type (the 2-talker masker resulted in higher SRTs than the speech-shaped noise masker). There was no significant interaction between group and masker, indicating that the groups did not differ in amount of perceptual masking.

The statistical outcomes across all postsurgery analyses were in agreement with each other. The 2-talker masker resulted in significantly higher SRTs than the speech-shaped noise masker, but there was no significant interaction between group and masker, indicating that the groups did not differ in amount of perceptual masking.

The results of this study indicated no effect of OME on perceptual masking, provided that the masker was continuous. However, when the masker was gated, children with OME history often showed more perceptual masking than children with no significant history of OME. It is not clear why masker continuity was a critical fac-
tor with regard to the effect of OME. One possibility is that there is a ceiling effect associated with perceptual masking for the continuous maskers. A previous study indicated that even normal-hearing young listeners have relatively great perceptual masking in continuous noise compared with adults (approximately 6 dB of perceptual masking for children vs approximately 2 dB for adults). Because perceptual masking is already relatively high in children, it may be difficult to show any additional effects related to OME.

It may be easier to demonstrate an effect of OME on perceptual masking in the gated condition because children with no history of OME show little perceptual masking here. Hall et al. found that both adults and children with normal-hearing history show a release from perceptual masking when the masker was gated (perceptual masking in a gated masker was eliminated in adults, and reduced to less than 2 dB in normal-hearing children). The present investigation indicates that perceptual masking often remains high in children with OME.

Table 2. Analysis of Covariance Results for the Continuous and Gated Maskers

<table>
<thead>
<tr>
<th>Testing Interval</th>
<th>Presurgery</th>
<th>1 mo</th>
<th>4-6 mo</th>
<th>12-13 mo</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Continuous Masker</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group</td>
<td>$F_{1,75} = 14.3; P &lt; .001$</td>
<td>$F_{1,75} = 0.26; P = .61$</td>
<td>$F_{1,59} = 0.5; P = .48$</td>
<td>$F_{1,59} = 1.2; P = .29$</td>
</tr>
<tr>
<td>Masker</td>
<td>$F_{1,75} = 101.4; P &lt; .001^*$</td>
<td>$F_{1,75} = 83.4; P &lt; .001^*$</td>
<td>$F_{1,59} = 65.8; P &lt; .001^*$</td>
<td>$F_{1,59} = 67.0; P &lt; .001^*$</td>
</tr>
<tr>
<td>Age</td>
<td>$F_{1,75} = 6.0; P = .002^*$</td>
<td>$F_{1,75} = 8.4; P &lt; .005^*$</td>
<td>$F_{1,59} = 4.4; P = .04^*$</td>
<td>$F_{1,59} = 10.9; P = .002^*$</td>
</tr>
<tr>
<td>Group × masker</td>
<td>$F_{1,75} = 0.1; P = .75$</td>
<td>$F_{1,75} = 0.1; P = .71$</td>
<td>$F_{1,59} = 0.8; P = .37$</td>
<td>$F_{1,59} = 0.22; P = .64$</td>
</tr>
<tr>
<td><strong>Gated Masker</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group</td>
<td>$F_{1,57} = 2.3; P = .13$</td>
<td>$F_{1,57} = 2.9; P = .10$</td>
<td>$F_{1,55} = 1.8; P = .19$</td>
<td>$F_{1,55} = 0.05; P = .83$</td>
</tr>
<tr>
<td>Masker</td>
<td>$F_{1,57} = 17.2; P &lt; .001^*$</td>
<td>$F_{1,57} = 28.1; P &lt; .001^*$</td>
<td>$F_{1,55} = 16.8; P &lt; .001^*$</td>
<td>$F_{1,55} = 5.3; P = .02^*$</td>
</tr>
<tr>
<td>Age</td>
<td>$F_{1,57} = 6.0; P = .02^*$</td>
<td>$F_{1,57} = 8.2; P &lt; .005^*$</td>
<td>$F_{1,55} = 8.3; P = .006^*$</td>
<td>$F_{1,55} = 7.4; P = .009^*$</td>
</tr>
<tr>
<td>Group × masker</td>
<td>$F_{1,57} = 7.1; P = .01^*$</td>
<td>$F_{1,57} = 7.5; P = .002^*$</td>
<td>$F_{1,55} = 5.1; P = .03^*$</td>
<td>$F_{1,55} = 0.26; P = .61$</td>
</tr>
</tbody>
</table>

*Effects were significant beyond a chance criterion of .05.*
history for the gated masker condition, at least up to approximately 6 months after tympanostomy tube placement. Hall et al\textsuperscript{14} proposed 2 explanations for the release from perceptual masking due to masker gating. One possibility considered was that a continuous speech masker results in an obligatory analysis of the running content. This analysis decreases the overall availability of the speech decoding resources for target word recognition, thereby worsening the recognition threshold. It was speculated that the gated masker was less effective because the available speech decoding resources were not preemptively engaged by an ongoing competing masker. The other possibility considered was that the gated masker may provide a salient auditory cue for when to listen for the target speech sound, reducing temporal uncertainty. Although there is a visual cue for when to listen in the continuous masker condition, this cue may not be as effective as the auditory cue provided by masker gating.

Regardless of the reason for the release from perceptual masking in the gated condition, it is interesting that the perceptual masking is relatively great in this condition for the children with OME history. This may reflect a relative inefficiency in separating simultaneously present speech streams and identifying the target speech element. This account would be consistent with previous psychoacoustical findings suggesting that children with OME history had difficulty segregating spectral components on the basis of coherent patterns of across-frequency amplitude fluctuation.\textsuperscript{12}

**RELATION TO PREVIOUS SPEECH STUDIES**

At least 2 previous OME studies\textsuperscript{1,2} have examined speech perception in the presence of a single-talker speech masker. Both of these studies found that children with OME history required higher signal-masker ratios than children without OME history. Unfortunately, it is not possible to interpret these results in terms of perceptual masking, per se. One reason for this is that there was no accompanying condition that incorporated a noise masker for comparison. Another reason is that a single-talker masker was used, and speech perception performance is better for a single-talker masker than for a noise masker.\textsuperscript{13} It was suggested\textsuperscript{13} that the single-talker masker is relatively ineffective because it possesses abundant temporal “windows” during which the target word can be processed. It is therefore possible that several factors could have contributed to the poor performance of the OME-history listeners in the single-talker case, including poor temporal analysis, factors related to perceptual masking, or general inefficiency in processing speech information (that would have been observed even with a noise masker).

Another result of the present study that deserves further comment is that, except for the presurgery case in continuous masking, listeners with OME history showed no deficit for spondee recognition in speech-shaped noise. Because such a speech recognition deficit was absent even in the first postsurgery test, it is parsimonious to attribute the poor presurgery result to audiibility issues related to threshold sensitivity. At first glance, a finding of no effect of OME on speech understanding in noise would appear contrary to the findings of Schilder et al.\textsuperscript{3} Studying children with OME history (but with normal audiograms at the time of testing), Schilder et al\textsuperscript{3} found poor speech recognition in speech-shaped noise. It is possible that the difference in findings between studies is related to the nature of the speech material. An important difference between studies is that the present study used spondee target words whereas Schilder et al used monosyllabic words. Because monosyllabic words are less redundant than spondees, the material used by Schilder et al required more sophisticated auditory analysis for correct identification. Another important difference is that whereas the present study used a closed response set, Schilder et al used an open set. It is therefore possible that effects of OME on speech understanding are more readily demonstrated with nonredundant material presented in an open set context.

As noted above, some of the deficits in speech understanding occurred only in the presurgery case, where audiometric hearing loss existed. Although these results do not indicate an effect of OME that endures beyond resolution of hearing loss, they do suggest that children with OME and accompanying hearing loss are at a disadvantage for speech understanding in noise. This is relevant to the broader question of disadvantages faced by children with OME and to issues related to the importance of treatment of the condition. The present results support the common sense conclusion that children having OME with hearing loss are at a disadvantage for understanding speech in noise.

**RELEVANCE TO SPEECH PROCESSING IN REAL ENVIRONMENTS**

Some of the issues considered above invite speculation about the relation between OME history and the processing of speech in real environments. Although all of the conditions tested here were unnatural, the continuous masker case is more similar than the gated masker case to real-life situations, where the listener must recognize a target speaker in the context of simultaneously present ongoing competing speech streams. In the continuous masker case, the children with OME history showed no greater perceptual masking than the children without OME history. This may suggest that although children, in general, have more difficulty than adults in processing a speech target in the presence of competing speech, children with OME history may not have any additional disadvantage. A caveat to this interpretation is that any such OME-related impairment may depend strongly upon the speech material being processed. For example, it is possible that an effect of OME on separating ongoing, simultaneously present speech messages might be present for relatively nonredundant, low-predictability speech material. Insight into this issue cannot be obtained from the present study because we used the classic approach of investigating perceptual masking, using material of relatively high predictability and redundancy (spondees).

Submitted for publication September 24, 2002; final revision received February 11, 2003; accepted February 12, 2003.
This research was supported by grant R01 00397 from the National Institutes of Health, National Institute on Deafness and Other Communication Disorders, Bethesda, Md.

Corresponding author: Joseph W. Hall, PhD, Division of Otolaryngology, The University of North Carolina at Chapel Hill, Chapel Hill, NC 27599.

REFERENCES