

Effect of masker type and age on speech intelligibility and spatial release from masking in children and adults

Patti M. Johnstone and Ruth Y. Litovsky^{a)}

Waisman Center, University of Wisconsin-Madison, Binaural Hearing & Speech Laboratory,
1500 Highland Avenue, Room 523, Madison, Wisconsin 53706

(Received 12 October 2004; revised 24 May 2006; accepted 19 June 2006)

Speech recognition in noisy environments improves when the speech signal is spatially separated from the interfering sound. This effect, known as spatial release from masking (SRM), was recently shown in young children. The present study compared SRM in children of ages 5–7 with adults for interferers introducing energetic, informational, and/or linguistic components. Three types of interferers were used: speech, reversed speech, and modulated white noise. Two female voices with different long-term spectra were also used. Speech reception thresholds (SRTs) were compared for: Quiet (target 0° front, no interferer), Front (target and interferer both 0° front), and Right (interferer 90° right, target 0° front). Children had higher SRTs and greater masking than adults. When spatial cues were not available, adults, but not children, were able to use differences in interferer type to separate the target from the interferer. Both children and adults showed SRM. Children, unlike adults, demonstrated large amounts of SRM for a time-reversed speech interferer. In conclusion, masking and SRM vary with the type of interfering sound, and this variation interacts with age; SRM may not depend on the spectral peculiarities of a particular type of voice when the target speech and interfering speech are different sex talkers. © 2006 Acoustical Society of America.
[DOI: 10.1121/1.2225416]

PACS number(s): 43.66.Lj, 43.71.Ft, 43.71.Gv, 43.66.Pn [GDK]

Pages: 2177–2189

I. INTRODUCTION

In a “cocktail party” environment (Cherry, 1953), a listener must be able to follow a specific conversation and ignore all the other interfering voices and sounds. Considerable research has been done to study how “cocktail party” environments affect adult listeners (for review, see Bronkhorst, 2000), however, very few studies have examined how children perform in a comparable situation. This area of research is rather important, given that children often find themselves having to hear and learn in very challenging acoustic settings.

The ability of children to extract information from auditory signals has been compared with that of adults in other measures. Children as young as 4 years of age demonstrate adult-like performance on frequency resolution tasks (Hall and Grose, 1994) and by 5 years of age minimum audible angle thresholds for simple sound configurations are adult-like (Litovsky, 1997). Neurophysiological studies support behavioral data in that children demonstrate adult-like frequency-specific maturation of the auditory periphery within the first year of life (Eggermont *et al.*, 1996; Ponton *et al.*, 1992).

In contrast, maturation is protracted past 10 years of age, for some temporal resolution tasks, thought to reflect central auditory processes (Hall and Grose, 1994; Hartley *et al.*, 2000). Temporal resolution abilities are likely important for sound source segregation in “cocktail party” environments. Other abilities that might similarly be important are

also not fully developed in young children. Examples include echo suppression (e.g., Morrongiello *et al.*, 1984; Litovsky, 1997; for a review, see Litovsky and Ashmead, 1997) and detection of movement of a fused auditory image (Cranford *et al.*, 1993). Similarly, the ability to detect signals in a fluctuating noise background is reduced in children when compared to adults (Grose *et al.*, 1993; Veloso *et al.*, 1990).

Studies on the “cocktail party” in adults, in which spectral cues vary, have shown that speech intelligibility depends on the number and type of interferers present (Culling *et al.*, 2004; Hawley *et al.*, 2004). When a single interfering sound is present, adults experience less masking in the presence of speech-based sounds such as speech and time-reversed speech than they do with modulated noise. When multiple interferers are present, however, speech reception thresholds (SRTs) with speech or time-reversed speech interferers become worse than for modulated noise. The trend toward increased masking with multiple speech interferers does not continue beyond two interferers. As the numbers of speech interferers is increased beyond two or more the difference between speech interferers and noise interferers decreases.

Another important factor is the advantage for speech intelligibility typically observed when the interfering sounds are spatially separated from the target, known as spatial release from masking (SRM) (Arbogast *et al.*, 2002; Bronkhorst and Plomp, 1988, 1992; Culling *et al.*, 2004; Dirks and Wilson, 1969; Freyman *et al.*, 1999; Hawley *et al.*, 1999; Hawley *et al.*, 2004; Peissig and Kollmeier, 1997; Plomp and Mimpen, 1981). In adults, SRM can be as large as 12 dB in the free field when multiple linguistically relevant interferers (such as speech or time-reversed speech) are present. SRM is smaller for adults in the presence of multiple

^{a)}Author to whom correspondence should be addressed; electronic mail: litovsky@waisman.wisc.edu

noise-type interferers that do not contain linguistic content or context (Bronkhorst, 2000; Hawley *et al.*, 2004). This suggests that when informational masking occurs, spatial cues become particularly important for source segregation (e.g., Bronkhorst, 2000; Culling *et al.*, 2004). The role of informational masking in studies that focus on free field spatial segregation remains to be better understood. Certainly for children, this remains an unexplored area.

When considering how children negotiate their acoustic space some important factors to bear in mind include: what *types* of sounds are interfering with the signal, where those interfering sounds are positioned relative to the target source, and *how many* interfering sounds are present. While studying these questions in 4–7 years-old children and adults Litovsky (2005) found that masking is greater with interfering sounds that do not contain linguistic content (e.g., modulated noise) than for speech, but SRM is similar for the two interferer types. The modulated noise (MN) interferer differs from speech in two ways. First, although amplitude fluctuations are similar to those found in speech, the spectrum of MN does not vary in time, while speech contains both amplitude and frequency fluctuations. Second, MN has no linguistic content or context. From that work, the relative effects of variation in spectral overlap and the presence or lack of linguistic content/context were difficult to tease apart. The current study therefore extended previous work to examine masking and SRM with time-reversed speech. Reversed speech carries linguistic *context* (listeners perceive it as being speech-like) but does not carry linguistic *content* (listeners cannot understand what is being said).

Like speech, ongoing frequency variations result in variable overlap between the first and second formants in the target and interfering sounds. It is also important to consider whether the amount of masking that is produced by a specific voice influences the size of SRM. Specifically, female voices can be quite variable with regard to the spectra (Stelmachowicz *et al.*, 1993; Hazan and Markahm, 2004), such that some can produce significantly more masking than others.

In the present study, three issues were examined. First, in one set of subjects three types of interfering sounds (modulated noise, speech, and time-reversed speech) were used in order to understand how stimuli that have linguistic content and/or context, or neither, affect masking and SRM in children and adults. Second, in two new groups of subjects the issue of female voice type was teased apart by comparing masking and SRM for two female voices. Both voices are pleasant to the ear and intelligible; while one has reduced energy at 1–3 kHz (female A), the other has a more “average” spectral shape and hence greater overlap in energy with the target (female B). Third, the effect of task difficulty on masking and SRM in adult listeners was examined by comparing the performance of the first group of adults using a 25-AFC task with published data from a group of adults using a 4-AFC task.

II. METHODS

A. Listeners

Twenty volunteer children in kindergarten and first grade were recruited from the Madison area public schools.

Twenty adult volunteers were recruited from the University of Wisconsin-Madison student and staff population. All of the listeners were native speakers of English and had normal hearing sensitivity, as indicated by pure-tone, air-conduction thresholds of 20 dB HL or less (ANSI, 1989) at octave frequencies between 250 and 8000 Hz. No asymmetry in hearing exceeded 10 dB HL between the two ears. Since middle ear problems are common in young children, tympanometry was performed before each visit using a screening tympanometer calibrated to ANSI specifications (ANSI, 1987). Children were included in the study only when their peak-compensated static admittance was normal.

The children comprised two groups with ten subjects in each group: Group 1A (ages 5.0–6.11) were tested using three different interfering stimuli (MN, female-A speech, female-A reversed speech), and completed testing in two 1-h sessions. Group 1B (ages 6.0–7.0) were tested using only the female-B speech interferer during a single session. The adults also comprised two groups with ten subjects in each group: Groups 2A (ages 18–42) and 2B (ages 18–30) were tested on parallel conditions to those of the corresponding children groups, and each subject completed testing in a single session. The children groups did not entirely overlap in age, but this is not a concern since Litovsky (2005) showed that masking and SRM do not vary under these conditions between 4.5 and 7 years of age.

B. Stimuli

The target stimuli consisted of a closed set of twenty-five, two-syllable children’s spondees recorded with a male voice and obtained from Auditec. The root-mean-square values of all words were equalized.

Four types of interfering stimuli were used. (1) forward speech spoken by *female A* with relatively little energy at high frequencies; (2) modulated white noise (MN); (3) time-reversed speech spoken by *female A*; (4) forward speech spoken by *female B* with greater energy at high frequencies, resembling the “average” spectra from previously published reports (e.g., Stelmachowicz *et al.*, 1993). Figure 1 shows the long-term spectra of the two female voices and the MN, each beside the target voice; these were obtained by taking the FFT of the signals in 1/3 octave bands.

The content of the speech interferers for both female voices consisted of digitized sentences from the Harvard IEEE list (Rothausser *et al.*, 1969). Examples of sentences include: “Tea served from the brown jug is tasty” or “A dash of pepper spoils beef stew.” To generate the MN interferer, the speech envelope was extracted from speech interferers created using the female-A voice and was used to modulate white noise tokens, giving the same coarse temporal structure as the speech. The envelope of running speech was extracted using a method similar to that described by Festen and Plomp (1990), in which a rectified version of the wave form is low-pass filtered. A first-order Butterworth low-pass filter was used with a 3-dB cut-off at 40 Hz. The time-reversed interferers consisted of the speech interferers, but simply reversed in time end to end (e.g., Hawley *et al.*, 2004). Though they shared the same temporal-spectral struc-

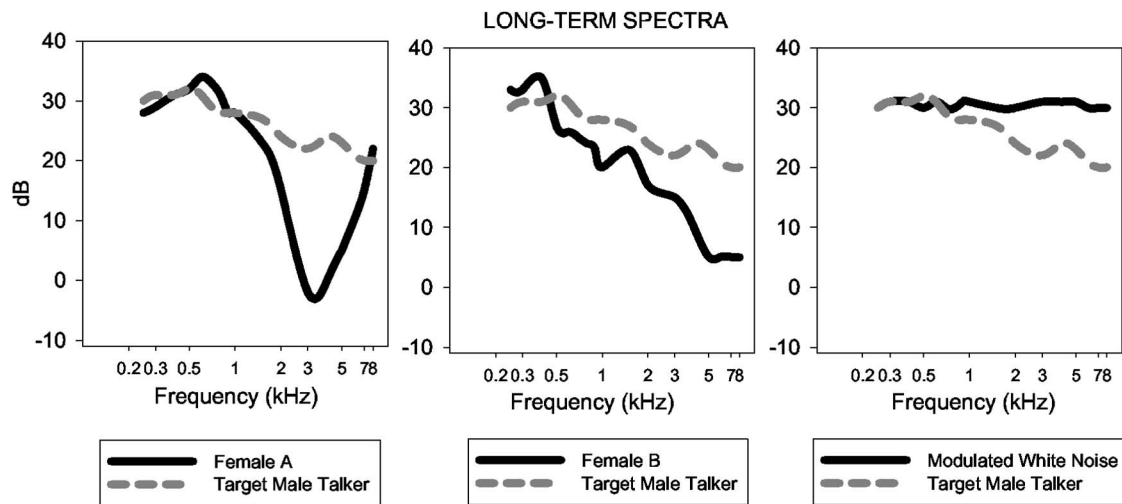


FIG. 1. The long-term speech spectra of two female talkers and MN are each shown beside the target male talker stimulus.

ture of the speech interferers using the voice of female A, the reversed speech interferers were unintelligible.

Testing was conducted in a standard IAC sound-proof booth with an inside dimension of 2.75 m × 3.25 m. Targets and interferers were presented via two separate loudspeakers (Cambridge Soundworks, Center/Surround IV) positioned at a distance of 1.5 m from the listener’s head.

C. Design and procedure

The three basic conditions were: Quiet (target 0° front, no interferer), Front (target and interferer both 0° front), and Right (interferer 90° right, target 0° front). In group 1A, children visited the lab twice. During their first visit testing included three conditions: Quiet, reversed speech-Front and reversed speech -Right. During the second visit testing included five conditions: Quiet, MN-Front, MN-Right, Speech-Front, Speech-Right. Children in group 1B were tested on three conditions: Quiet, Speech-Front, Speech-Right. During each session, the order of conditions was randomized for each child, and testing was conducted once per condition. Adults in group 2A completed two repetitions of each condition within a single visit, and adults in group 2B were tested once on each condition; order of presentation was always randomized.

Prior to testing, each child participated in a brief familiarization task to determine if s/he could readily identify every target word. The experiment was designed to measure the effect of speech and noise interferers on word recognition, rather than vocabulary. During the familiarization task the child was asked to identify the pictured spondees. For each participant that was recruited, within minutes, it was clear that the child was comfortable and familiar with every target word that could be used during the experiment.

All listeners sat at a small table facing a computer screen that was positioned under the loudspeaker at 0°.

In conditions containing interfering sounds, the interferer was turned on first followed by the target presentation, and continued after the target was turned off for approximately 1–2 s. Subjects were instructed to ignore the female voice and to listen carefully to the male voice.

The task for children consisted of a 4-alternative-forced-choice procedure (Litovsky, 2004; 2005). On every trial a word from the Children’s Spondee List, spoken by a male talker, was chosen randomly from a closed set of 25 targets. The randomization process ensured that for every subject, on average, all 25 words were selected an equal number of times. The target word was preceded by a leading phrase “Ready? Point to the....” also spoken by a male talker. Each child was then asked to identify a picture matching that word from an array of four pictures that appeared together on the computer screen, only one of which matched the target. Correct responses were followed by positive feedback (animation); incorrect responses were followed by “negative” feedback from the computer such as “let’s try another one” or “that must have been difficult.”

The task for adults in group 2A consisted of a 25-alternative-forced-choice (25-AFC) procedure. Subjects used the computer mouse to select the written word from a list of 25 words that appeared on the computer screen, only one of which matched the target. No feedback regarding performance was provided. The 25-AFC task in the current experiment was invoked as a means of equating task difficulty as much as possible between adults and children. In the previous study, adults tested on the 4-AFC task (Litovsky, 2005) had very low SRTs compared with children. In addition, there was little or no SRM on some conditions, suggesting that the ease with which the task was performed may have eliminated some of the most interesting effects. In the present study, an effort was made to balance age-dependent difficulty of the task by increasing the number of alternatives in the forced choice paradigm.

D. SRT estimation

At the start of each condition measurement, the level of the target was initially 60 dB SPL. When interferers were present (Front or Right conditions), the interferer level was fixed at 60 dB SPL. SRTs were estimated using a method described by Litovsky (2004, 2005). An adaptive tracking method was used to vary the level of the target signal, such that correct responses result in level decrement and incorrect

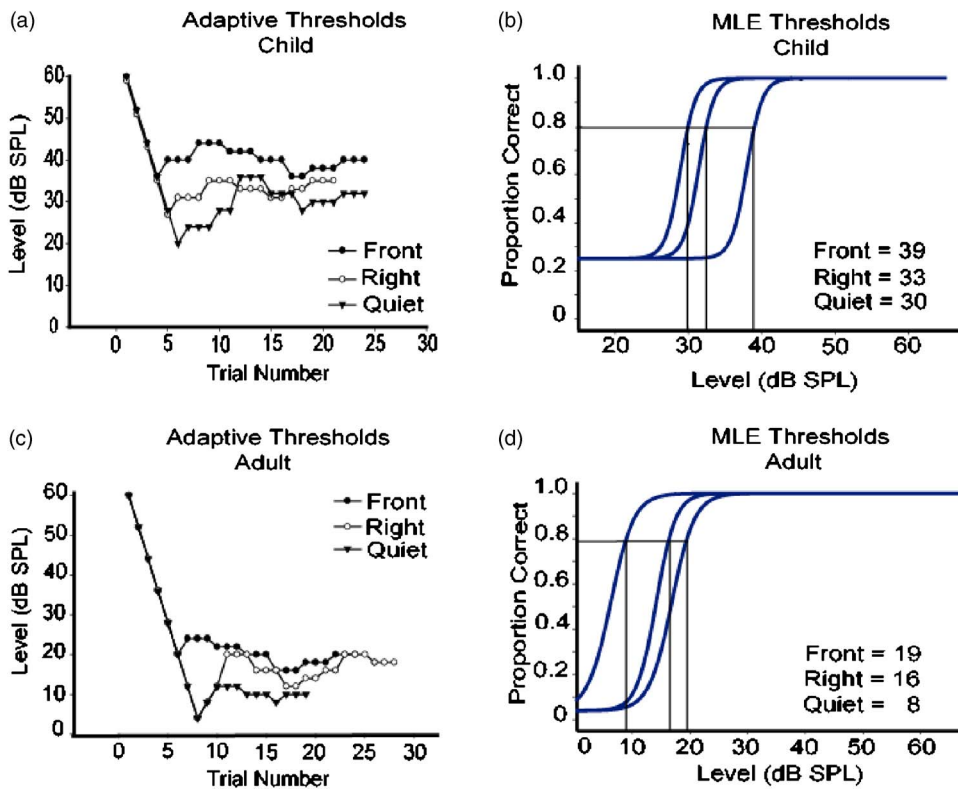


FIG. 2. Examples of adaptive tracks (3-down-1-up) for one child (A) and one adult (C) are shown. Interferer level remained fixed at 60 dB SPL. A sigmoidal function [(B), (D)] was fit to the raw data from each adaptive track, and a maximum likelihood estimate (MLE) procedure was used to estimate threshold at a performance level of 80% correct. The legend reports the associated estimates.

responses result in level increment. The algorithm includes the following rules: (1) Level is initially reduced in steps of 8 dB, until the first incorrect response. (2) Following the first incorrect response a 3-down/1-up rule is used, whereby level is decremented following three consecutive correct responses and level is incremented following a single incorrect response. (3) Following each reversal the step size is halved. (4) The minimum step size is 2 dB. (5) A step size that has been used twice in a row in the same direction is doubled. For instance if the level was decreased from 40 to 36 (step=4) and then again from 36 to 32 (step=4), continued decrease in level would result in the next level being 24 (step=8). (6) After three consecutive incorrect responses a “probe” trial is presented at the original level of 60 dB. If the probe results in a correct response the algorithm resumes at the last trial before the probe was presented. If more than three consecutive probes are required, testing is terminated and the subject’s data are not included in the final sample. (7) Testing is terminated following four reversals [e.g., Figs. 2(a) and 2(c)].

SRTs were estimated using a constrained maximum-likelihood method of parameter estimation (MLE) outlined by Wichmann and Hill (2001a; 2001b). All the data from each adaptive track were fit to a logistic function and the inverse of the function at a specific probability level was taken. Slopes were calculated by taking the derivative of the function with respect to threshold. Psychometric functions for the children’s data, which were collected with a 4-AFC task, were set to a lower bound level of 0.25, which was the level of chance performance. Given that an adaptive 3-down/1-up procedure was used, threshold corresponded to the

stimulus value point on the psychometric function where performance was approximately 79.4% correct (as estimated by Levitt, 1971).

It is important to note that biased estimates of threshold can be introduced by the sampling scheme used and lapses in listener attention. The upper bound of the psychometric function was constrained within a narrow range (0.05), as suggested by Wichmann and Hill (2001b), who demonstrated that bias associated with attention lapses was overcome by introducing a highly constrained parameter to control the upper bound of the psychometric function. As the authors suggest, under some circumstances, bias introduced by sampling scheme may be more problematic to avoid even when a hundred trials are obtained per level visited.

SRTs obtained using MLE were compared with SRTs calculated from the last three reversals in each experimental run. A repeated measures *t*-test revealed no statistical difference between the two threshold estimates [$t(179)=1.832$, $p=0.0686$], two tailed]. Although the MLE procedure produces similar group mean thresholds, its advantage is that group variances are typically smaller. The MLE approach offers an advantage especially when few reversals can be obtained, such as when working with young children, in which case a single trial can have a disproportionate weight. The MLE approach reduces the effect of individual trials by placing greater weight on levels at which the total number of trials is largest.

E. Data analysis

Average SRTs and standard deviations for children and adults for all conditions (Quiet, Front, Right) and interferer types were computed. Since the children in group 1A visited

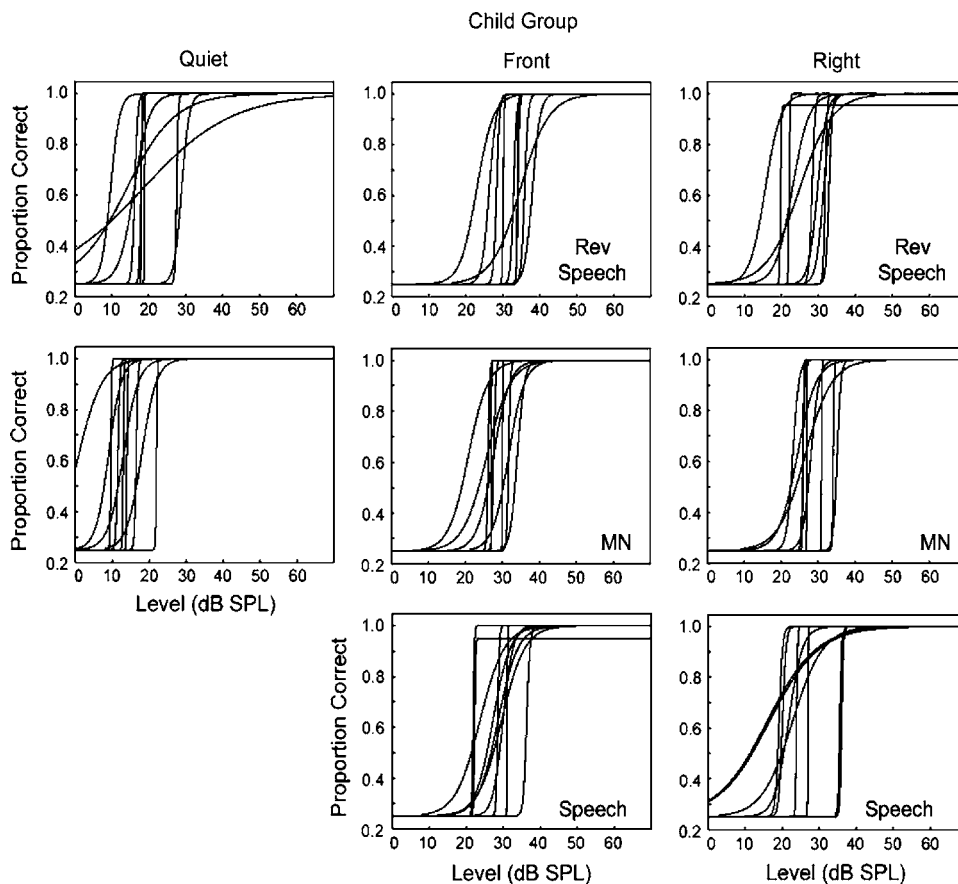


FIG. 3. Sigmoidal functions are shown for each individual measurement in the child group. Each panel shows functions for one condition (Quiet, Front, Right) and one interferer type.

the lab twice, the possibility that a learning effect occurred was considered upon comparison of thresholds obtained during the two visits. It appeared that the children's SRTs were generally higher for all conditions tested on the first visit, including Quiet, compared with those tested on the second visit. Nevertheless, the two Quiet thresholds were not statistically different ($p > 0.05$).

The higher thresholds obtained during the first visit would not have posed a problem if the type of interferer encountered during the first visit was randomized among the children in group 1A. This approach was not taken because the study was initially designed for the purpose of measuring performance with the reversed speech only. When the results appeared to be quite different between the child and adult groups a post-hoc decision was made to retest the same children using the speech and MN interferers. The children returned to the lab for a second visit to collect the speech and MN interferer data. This same step was not performed for adults in group 2A, because they were tested on all conditions during the initial visit.

In order to reduce the potential biasing effect that learning in a test-retest situation might create, all thresholds from the second visit were normalized by subtracting the average difference between the two visits for the Quiet condition (3.7 dB). It is important to note that, although the reduction in the Quiet condition may overestimate changes due to retesting that might occur in the Front and Right conditions, it was the only condition that was retested during the second visit and could be directly compared for threshold changes. While future studies should examine test-retest issues across

conditions more systematically, for the purpose of the present study, the decision to normalize using Quiet thresholds is the most conservative approach, effectively reducing the age difference in the reversed speech conditions. All statistical analyses were performed using the normalized data.

III. RESULTS

Results are first discussed for groups 1A and 2A, who were tested on three types of interfering sounds (MN, speech, and time-reversed speech). Second, results from the speech conditions are compared for the groups that were tested with two different female voices (female A, female B). Finally, results with adult subjects (group 2A) using the 25-AFC procedure are compared to published data from adult subjects using the 4-AFC procedure (Litovsky, 2005). The voice of female A was used as the speech interferer for both groups.

A. Estimating SRTs and goodness of fit

Because results presented here essentially consist of SRTs that are based on estimates of psychometric functions, individual fits to the data are shown in Figs. 3 and 4, for groups 1A and 2A. Quiet psychometric functions are generally shifted toward lower signal levels relative to data collected during the Front and Right conditions. Also noticeable is that the functions in adult subjects (Fig. 4) ascend at lower levels and appear to be less variable within each condition than the children's functions (Fig. 3). The psychometric functions also clearly vary in steepness. It is difficult to know

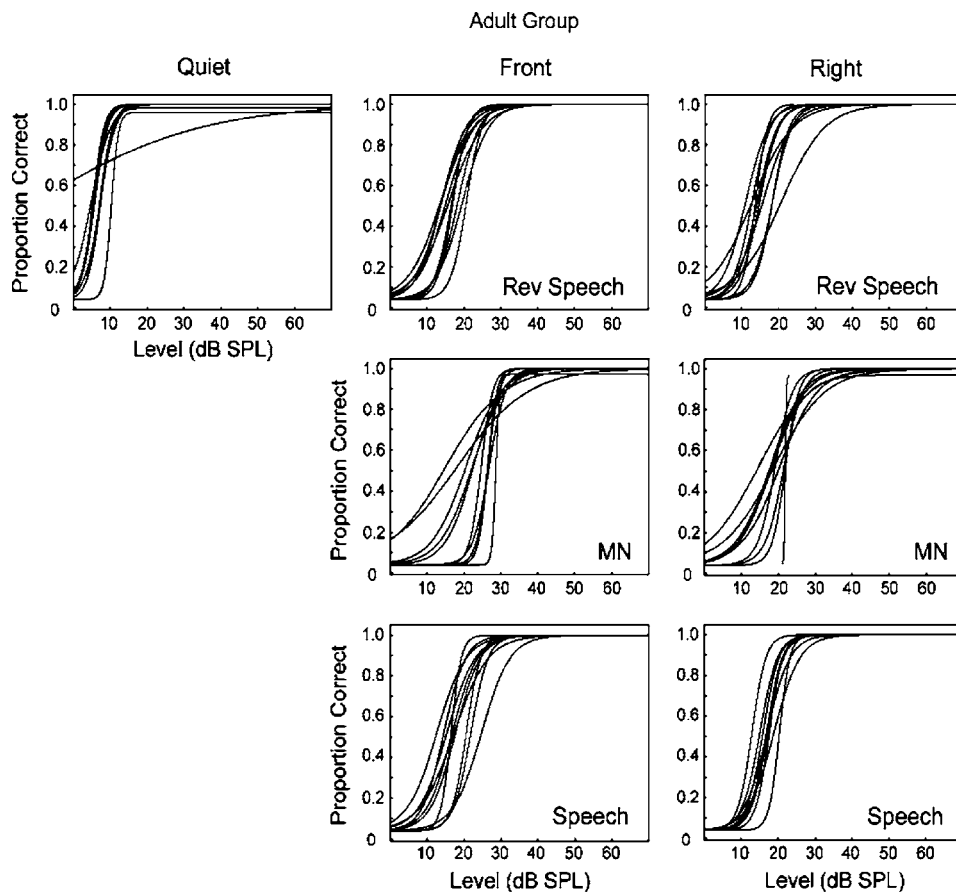


FIG. 4. Sigmoidal functions are shown for each individual measurement in the adult group. Each panel shows functions for one condition (Quiet, Front, Right) and one interferer type.

if the slope of the psychometric function is a true reflection of a sensory process or is due to averaging high variability in the underlying track (Leek *et al.*, 1991b). Individual tracks that “wandered” did result in shallow functions for both adults and children. The high variance occasionally seen in individual tracks obtained in this study did result in shallow psychometric functions as would be expected. Allen and Wightman (1995) attributed shallow psychometric functions obtained in their study to inattention. It is possible that subjects in our study experienced lapses in attention and that contributed to reduction in slopes of some of the psychometric functions.

The unusually steep functions obtained in this study were associated with dramatic changes in performance over a very small change in level. This pattern was seen more frequently with the children, whose performance often went from 100% correct to 0% correct over a very small decrement in level. It is possible that over a small change in level the target went from being unintelligible to being clearly intelligible. It is also possible that children are not particularly good at guessing the correct response when only a portion of the target word is heard and understood.

As is often the case with populations of subjects whose performance can vary dramatically, an objective analysis of the goodness of fit of the psychometric functions is appropriate. This was determined using deviance as described by Wichmann and Hill (2001a, 2001b) and as applied to informational masking data by Lutfi *et al.*, (2003).¹ Slightly less than 9% of the fitted functions were estimated to have a likelihood of less than 5%. Only three of the psychometric

functions had estimated parameters that identified them as clear outliers. Two of these outliers were in the Quiet condition (one child and one adult) and one was for a child in the Right, MN condition.

In addition to measuring the deviance for each psychometric function, MLEs of confidence intervals were obtained for each threshold estimate following the procedure outlined by Lutfi *et al.*, (2003), and using the method devised by Wichmann and Hill (2001a, 2001b). A bootstrapping technique was used to estimate the sampling distributions of each threshold and determine confidence limits for the estimated distributions. Using proportion correct (assuming a binomial distribution) at each signal level from each fitted function, a simulated proportion correct at each signal level was randomly drawn and a logistic was fit to these values. This procedure was repeated 10 000 times to provide 10 000 estimates of the logistic function. Confidence limits were set at 0.025 and 0.975 so the points from the sampling distribution of the thresholds provided a 95% confidence interval. Individual thresholds fell within the confidence interval.

B. Effect of age on SRTs

In general, children had higher SRTs than adults. These results are consistent with previous research using the same methods and procedures (Litovsky, 2005). Figure 5 shows average SRTs (\pm s.d.) for children and adults for the various conditions and interferer types. For both children and adults, SRTs were higher in the presence of a single interfering sound source compared with the Quiet condition. Overall the

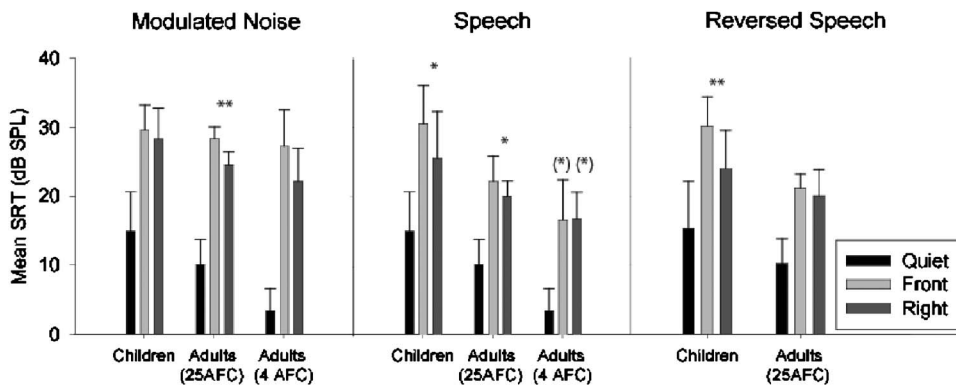


FIG. 5. Mean (\pm s.d.) SRTs are shown for children and adults, for the different types of interferers and conditions. Asterisks indicate significant difference between group means in the Front and Right conditions: * $p < 0.05$, ** $p < 0.01$. Tables I and II identify the other significant differences. Data using a 4-AFC procedure from Litovsky (2005) are also included for adults for MN and speech interferers. Asterisks in parentheses indicate significant differences between the 4-AFC procedure and the 25-AFC procedure.

increase was 6–15 dB for children and 10–17 dB for adults. A mixed, nested analysis of variance (ANOVA) was performed to test for age differences in SRTs. Two within-subjects variables [interferer-type (MN, speech, reversed speech) and condition (Quiet, Front, Right)] and a single between-subjects variable [age (child, adult)] were used. Results revealed no significant main effects, but a significant three-way interaction ($F[4, 72] = 3.097, p < 0.01$).

Tables I and II list significant differences obtained with post hoc Scheffé tests for group SRTs. Children had significantly higher SRTs than adults when the interferer was speech ($p < 0.01$) or reversed speech ($p < 0.001$) in the Front condition (at the same location as the target). In the Right condition, children had significantly higher SRTs than adults for all interferer types ($p < 0.05$). In the Quiet condition, there was no difference between SRTs of the two age groups.

Post hoc Scheffé tests of adult data also revealed that in the Front condition SRTs were higher (greater masking) for the MN interferer compared with speech ($p < 0.001$) or reversed speech ($p < 0.0001$). This pattern of results was also found in the Right condition ($p < 0.0005$ and < 0.0001 for speech and reversed speech, respectively). SRTs did not differ between the two speech-type interferers. Finally, SRTs were significantly higher in the Front than Right conditions, for the MN ($p < 0.01$) and speech ($p < 0.05$) interferers, but not for reversed speech. Children’s performance differed from adults in the Front condition; a lack of significant post hoc effects suggests similar amounts of masking for all interferer types. In contrast, in the Right condition children’s performance mirrored that of adults in that SRTs were significantly higher with the MN than with the speech ($p < 0.01$) or reversed speech ($p < 0.0001$) interferers.

One of the hallmarks of auditory development is indi-

vidual variability. While some children reach adult-like performance at relatively young ages, others do not. In addition, since the task focuses on listening to speech in the presence of maskers, the actual signal-to-noise ratio (SNR) at which SRTs are achieved is an important feature of the data that can then be used to compare with other masking studies. Therefore, Fig. 6 shows individual results on all conditions, as a function of the SNR at which SRTs were obtained. The extremely low SNR values (–30 to –40) reported here are likely due to several factors: the type of target used (simple spondee), the number of interferers (a single one), the difference in F0 between the target (male voice) and interferer (female voice), and the spectral characteristics of the interferer (female voice) used in our experiments. These findings are consistent with results obtained in a previous study using a similar test procedure (Litovsky, 2005).

C. Age effects and the ability to utilize interferer-dependent cues

Masking was quantified from the measured SRTs in two ways: Quiet SRTs were subtracted from Front (F-Q) or Right (R-Q) SRTs, resulting in values corresponding to “Front Masking” and “Right Masking.” Masking was evaluated using a mixed ANOVA with two within-subjects variables [interferer-type (MN, speech, reversed speech) and masking condition (F-Q, R-Q)] and one between-subjects variable [age (child, adult)].

TABLE II. Comparisons made between differences in mean SRTs for the different interferer types and the level of significance obtained using post hoc Scheffé test.

Condition	Group	SRT comparison	Significance
Front	Adults	MN > Speech	$p < 0.001$
Front	Adults	MN > Rev Sp	$p < 0.0001$
Front	Adults	Speech vs Rev Sp	Not significant
Right	Adults	MN > Speech	$p < 0.0005$
Right	Adults	MN > Rev Sp	$p < 0.0001$
Right	Adults	Speech vs Rev Sp	Not significant
Front	Children	MN vs Speech	Not significant
Front	Children	MN vs Rev Sp	Not significant
Front	Children	Speech vs Rev Sp	Not significant
Right	Children	MN > Speech	$p < 0.01$
Right	Children	MN > Rev Sp	$p < 0.0001$
Right	Children	Speech vs Rev Sp	Not significant

TABLE I. Comparisons made between differences in mean SRTs for groups 1A and 2A and the level of significance obtained using post hoc Scheffé test.

Condition	Interferer type	SRT comparison	Significance
Front	MN	Children > Adults	$p < 0.05$
Front	Speech	Children > Adults	$p < 0.01$
Front	Reversed speech	Children > Adults	$p < 0.001$
Right	MN	Children vs Adults	Not significant
Right	Speech	Children > Adults	$p < 0.05$
Right	Reversed speech	Children > Adults	$p < 0.05$
Quiet	No interferer	Children vs Adults	Not significant

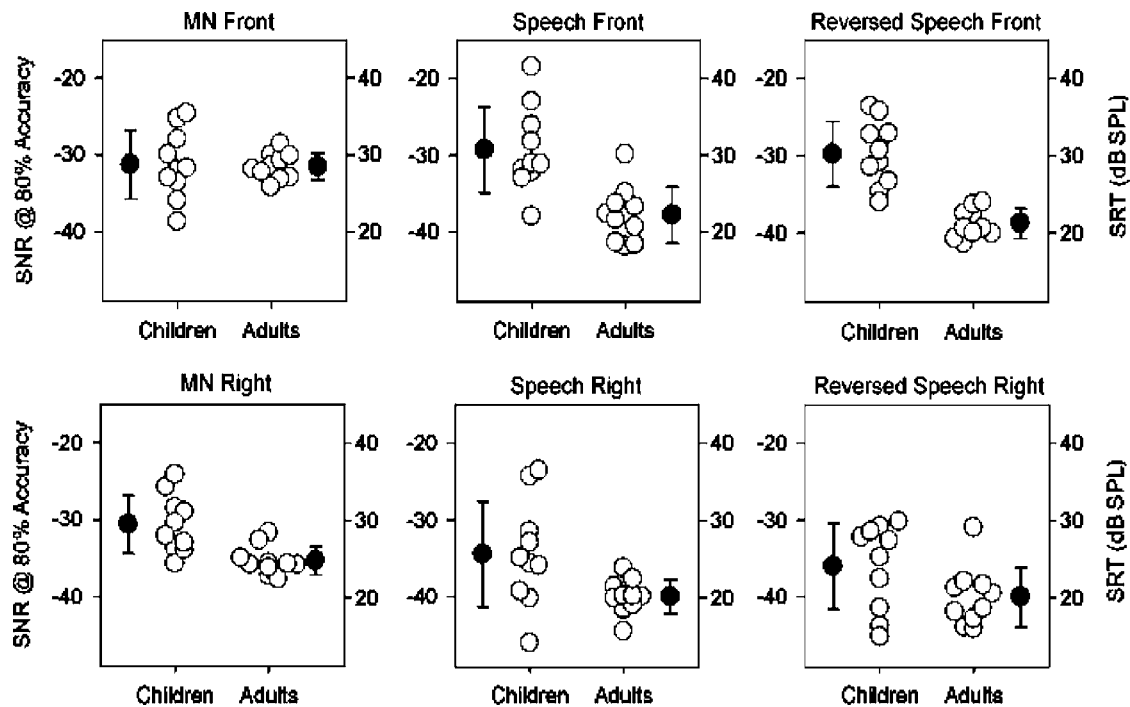


FIG. 6. Performance of children and adults is compared for the three types of interferers. Data are plotted such that either SNR at which performance reached 80% correct (left axis value) or SRT (right axis value) can be seen. Individual data (white symbols) and group means \pm s.d. (black symbols) are shown.

No significant main effects for masking were found. However, there was a significant 3-way interaction ($F[2,36]=5.645$, $p<0.01$). Post hoc Scheffé tests revealed that, in the adult group F-Q masking was significantly larger with the MN than speech ($p<0.0001$) or reversed speech ($p<0.0001$). Children, however, did not show any significant differences in the amount of F-Q masking across interferer types. Interesting age effects were found for F-Q masking. Children experienced significantly greater masking than adults with reversed speech ($p<0.01$), but significantly less masking than adults with MN; there were no age effects in Front masking for the speech interferer.

In the Right conditions, both children and adult groups showed more masking with MN than speech ($p<0.01$) or reversed speech ($p<0.01$). Children and adults did not differ significantly on the amounts of R-Q masking for any interferer type.

D. Spatial release from masking (SRM)

Spatial release from masking (SRM) is defined as the difference in SRT between the Front and Right conditions (F-R). A mixed ANOVA was applied to SRM values treating interferer type (MN, speech, reversed speech) as the within-subjects variable and age (child, adult) as the between-subjects variable. There were no significant main effects, but a significant two-way interaction was found ($F[2,36]=21.369$, $p<0.0005$). Post hoc Scheffé tests revealed that SRM was significantly *larger* in children than adults with reversed-speech ($p<0.01$), but significantly *smaller* in children than adults with MN ($p<0.05$). In addition, the amount of SRM for children was significantly less with MN compared with reversed speech ($p<0.01$) or speech ($p<0.05$), and there was no difference between speech and reversed

speech. In adults, SRM was similar for all three interferer types. These results suggest that SRM is interferer-type dependent for children but not for adults. In addition, SRM in children is best measured with a speech-type interferer rather than a noise-type interferer.

SRM is a measure that varies greatly across individuals. This can be seen in Fig. 7, where SRM becomes visible when R-Q values are plotted as a function of F-Q values for each subject, by condition. The line bisecting each panel represents unity, such that symbols falling along the diagonal indicate absence of SRM, while symbols appearing below the line indicate that SRM occurred. It is clear that although SRM occurred for the majority of subjects, especially in the speech-type conditions, there was great variability across individuals, and that the adult data are generally clustered more tightly than the children's data.

E. Effect of type of female voice

In an effort to understand how masking and SRM could be affected by variability in the amount of high-frequency energy present in female talkers, two female voices with very different voice spectra were used in the present study on some of the measures. Results were compared for the two groups of children (1A and 1B) and two groups of adults (2A and 2B). Plotted in Fig. 8(a) are values for F-Q and R-Q masking obtained with the two different female talkers, for children and adults. A mixed ANOVA was performed with one within-subjects variable [masking condition (F-Q, R-Q)] and two between-subjects variables [age (Child, Adult) and talker (Female A, Female B)]. Significant main effects were found for both masking condition [$F(1,36)=14.918$, $p<0.0001$] and talker [$F(1,36)=65.890$, $p<0.00001$], with no significant interactions. These results show that amount of

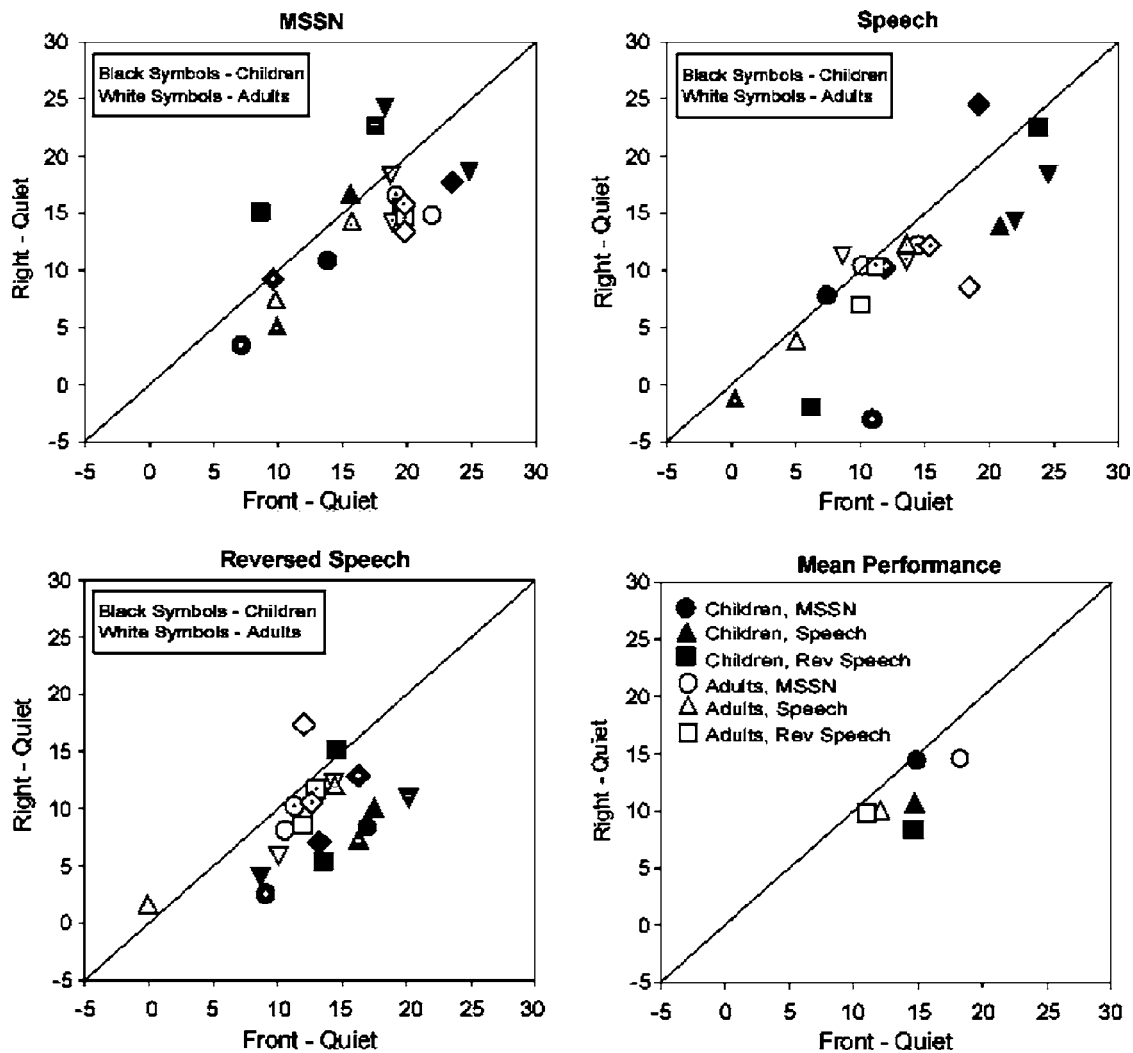


FIG. 7. Individual data points for children (black symbols) and adults (white symbols) are plotted for the R-Q condition as a function of values obtained in the F-Q condition. Each plot shows values for a different type of interferer. The plot on the bottom-right shows average values for these conditions, for the two age groups.

masking was higher for Female B than Female A, for both adults and children. In addition, masking was greater in the F-Q than R-Q condition, for both age groups, and both voices. These results are not surprising given that the voice of female B contains more spectral energy in the high frequencies. It is interesting to note that F-Q and R-Q masking increased equally, by about 11 to 12 dB, for both children and adults. The equal increase in masking in the Front and Right conditions [see Fig. 8(b)] suggests that the increase may be due to energetic masking.

SRM was evaluated with a two-way between-subjects ANOVA [age (Child, Adult); and talker (Female A, Female B)]. No significant main effects or interactions were found. The amount of SRM obtained using the two different female voices did not differ for either age group (see Fig. 9), suggesting that the effect being explored in this study remains constant regardless of the two types of voices used here.

F. Task difficulty in the forced choice paradigm for adults

The use of the 25-AFC task in the current experiment was a deliberate attempt to make the task more difficult for

the adults compared with the 4-AFC task. It was hypothesized that, by increasing the level of difficulty for the adults, SRTs would increase. Independent-samples *t*-tests were used to compare the adult data from the 4-AFC task (Front and Right, MN and speech interferers; Litovsky, 2005) with adult data obtained here using the 25-AFC. No statistically significant difference for the MN interferer was found in the Front or Right ($p > 0.05$). SRTs were significantly higher for the speech interferer with the 25-AFC task than the 4-AFC task in both Front [$t(18)$ one tailed = -2.586, $p < 0.05$] and Right [$t(18)$ one tailed = -2.280, $p < 0.05$]. Task difficulty seems to increase masking for a speech interferer, but not for a noise interferer. There was no effect of task difficulty on SRM, suggesting that with either method the adult and child data are comparable.

IV. DISCUSSION

This study investigated the ability of young children and adults to understand speech in the presence of three different types of interferers (MN, speech, and time-reversed speech), for two different female voices, and the extent to which lis-

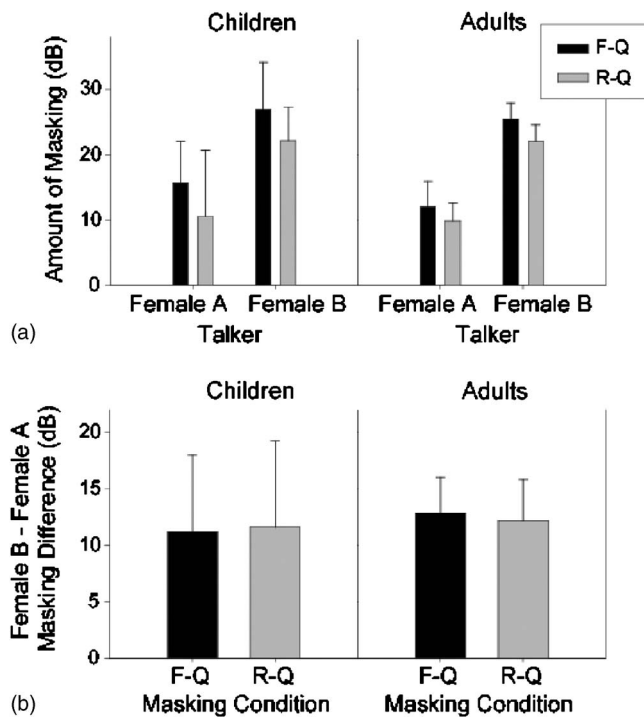


FIG. 8. (A) Mean (\pm s.d.) amount of masking in the F-Q and R-Q conditions, shown for children and adults. (B) Mean (\pm s.d.) increase in masking (F-Q and R-Q) for the Female B interferer compared with Female A. Values are shown for children and adults in dB.

teners benefit from spatial separation of the target and interferer. Effects of noise-based, language content-based and language-context based interference were compared. A child-friendly task was employed in which children could perform well enough to evaluate their performance under conditions that vary in difficulty. This approach requires some thought with regard to previous work conducted with adults, in which typically harder, open-set sentence materials are used as target stimuli.

A. Age effect on speech reception threshold

Results suggest that, compared with the Quiet condition, adding a single interferer leads to decreased speech intelligibility, as seen by elevated SRTs for both children and adults. Children's SRTs tended to exceed those of adults regardless of the type of interferer, location of the interferer, or level of task difficulty (4-AFC task for children versus 25-AFC task

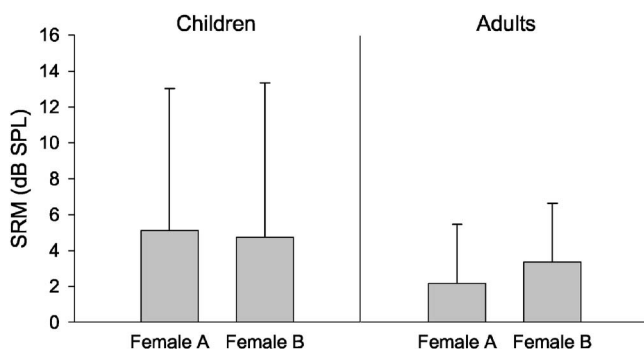


FIG. 9. Mean (\pm s.d.) amount of SRM for children and adults obtained using two different female voices.

for adults), see Fig. 4; this finding supports previous reports on speech recognition performance in adults and children in the presence of speech interferers (Litovsky, 2005; Papso and Blood, 1989) and speech-shaped noise (Hall *et al.*, 2002). These results can be explained by monaural masking differences, since young children usually experience higher thresholds than adults in the context of complex listening tasks (e.g., Grose *et al.*, 1993; Veloso *et al.*, 1990).

B. Interferer-dependent masking

The results suggest that the amount of masking varied by location, interferer type, and age. In adults, the MN interferer produced more masking than either speech or reversed speech. Effect of interferer type on masking measured here for adults is consistent with previous work of Hawley *et al.* (2004) in which masking with a single interferer was greater with MN than speech or reversed speech; the authors attributed this difference to an energetic masking effect. In the presence of a single speech or reversed-speech competing signal, listeners can take advantage of both the temporally modulated amplitude gaps and the spectral variation to hear out the target speech, but the lack of spectral modulation in MN results in greater energetic masking and greater overall SRTs.

Energetic masking, which is thought to occur when the signal is degraded by energetic spectral overlap from the interfering sounds, is presumed to occur in the peripheral auditory system (e.g., von Békésy, 1960; Green and Swets, 1974). This contrasts with informational masking which occurs when the signal is "lost" in a background of interference that shares similar, often nonoverlapping acoustic patterns with the signal (Watson *et al.*, 1976; Lutfi, 1990; Leek *et al.*, 1991a, b; Brungart, 2001; Kidd *et al.*, 2002; Arbogast *et al.*, 2002; Durlach *et al.*, 2003). Informational masking is thought to result from processes within the central auditory system that cannot be accounted for simply by peripheral mechanisms. The speech and reversed-speech interferers are good examples of maskers that produce a combination of energetic and informational masking, although the relative contributions of each effect are difficult to tease apart.

Masking also varied by location and type, especially in children: MN produced higher SRTs than speech or reversed speech when the target and interferers were spatially separated, but there was no effect of interferer type when the target and interferer were spatially coincident. In other words, in absence of spatial cues all interferers were equally difficult to ignore, but when spatial cues were available they were more useful with speech-based interferers than the noise-based interferer. Similar interactions between interferer type and location have been reported when spatial separation was induced using the precedence effect paradigm (Freyman *et al.*, 2001).

C. Age- and task-dependent masking

One explanation for the age difference in amount of Front masking produced by different interferers is that it may reflect maturation of frequency resolution. That is, children may be less able to use differences in the F0 between the

target and maskers in the speech and reversed-speech conditions, an ability that adults are known to be very good at when a single interferer is used (Festen and Plomp, 1990; Brungart *et al.*, 2001; Hawley *et al.*, 2004). This explanation is not compelling because basic abilities such as frequency resolution and tone detection in noise are adult-like by age 4 (e.g., Hall and Grose, 1994). More likely is the explanation that speech intelligibility in the presence of any masker, especially in absence of spatial cues, is more challenging for children than adults due to immature central auditory processes that are required for source segregation.

Our findings are consistent with reports on developmental changes in auditory attention in children (for reviews see Gomes *et al.*, 2000; Plude *et al.*, 1994). Younger children are more likely to process both relevant and irrelevant stimulus streams (Berman and Freidman, 1995; Doyle, 1973; Macoby, 1969), and seem to have difficulty selecting the appropriate stimulus channel and maintaining attentional focus over time (Berman and Friedman, 1995; and Macoby, 1969). In addition, older children (and adults) seem better able to engage in tasks and to employ useful strategies to solve problems that require selective auditory attention (Gutentag and Ornstein, 1990). It is possible that the underlying inefficient allocation of auditory attention often seen in younger children is directed not only by the physical characteristics of the stimuli but by the interests, motives, and cognitive strategies of the child perceiving them (Gomes *et al.*, 2000). This could likely affect performance during the most challenging conditions in the present study, such as when the target and interferer arise from similar locations.

Finally, it is important to note that the SNR at which SRTs are estimated were fairly low in the present study compared with related work. For instance, here the SNR for 80% accuracy was between -30 and -40 dB as compared to -10 dB reported by Hawley *et al.*, (2004). This can be explained by considering several methodological differences between the two studies. Here we used target and interferers that differed in gender (and fundamental frequency), while the other study used the same talker for both target and interferer. Second, we used a small corpus of targets in a closed-set forced-choice paradigm, whereas in the other study open-set sentences were employed. Third, we only had the single-interferer condition compared with single-, double-, and triple-interferer conditions in the Hawley *et al.* study. Additional interferers would have likely added more masking and reduced the threshold SNR. These issues are important to consider in future developmental studies. Finally, the spectral characteristics of the interferer (female voice) used in our experiments reduced the amount of masking experienced by listeners.

D. Spatial release from masking (SRM)

Results in the current study suggest that both children and adults experience SRM. Adults showed slightly larger SRM here than in the Litovsky (2005) study, likely due to the increased difficulty of the 25-AFC task compared with the 4-AFC task in the previous study. In general though, the average amount of SRM reported here for adults (about

2 dB) is at the lower range of previous reports of 2–4 dB for a single interferer (Hawley *et al.*, 1999, 2004). The small amount of SRM reported here may be attributed to several methodological differences. First, another level of task difficulty is introduced when open-set sentences are used (e.g., Hawley *et al.*, 2004) compared with the closed-set spondees, even in a 25-AFC task. Second, different gender voices were used for the target (male) and interferers (female-based) in the current study. Previous research has shown that adults can experience a 4–5 dB increase in SRM when the target and interferer are the same gender voices versus different gender voices and with a more difficult task (e.g., Brungart and Simpson, 2002a). Finally, a single interferer was used in the current study. Other research has shown that when multiple linguistically relevant interferers are present the amount of SRM also increases (Hawley *et al.*, 2004). One of the greatest benefits of SRM may be the improvement in speech intelligibility afforded under conditions in which the listening environments contain the most challenging content and/or contexts.

The children in this experiment demonstrate greater amounts of SRM with interferers that have speech-like properties (an informational masker) compared with the interferer that has the same long-term spectrum as the target (an energetic masker). In addition, children showed largest SRM for the time-reversed speech interferer (6.7 dB) compared with speech (3.4 dB) or MN (0.5 dB). We hypothesize that the novelty of a reversed speech signal affected how children allocated their attention. Anecdotally, one child asked the experimenter what language the lady was speaking. It may be that the children gave the reversed-speech interferer special attention because it sounded like a real, albeit foreign language to them. The adults, on the other hand, having had more extensive real-world knowledge would either not have mistaken it for a foreign language, or would have treated it like any other speech-like sound.

As an interferer, time-reversed speech is unusual. It has no semantic content, and reversal of speech in the time domain renders discrimination difficult if not impossible (Ramos *et al.*, 2000), in part because the reversal dramatically alters the onset and offset patterns for voicing and the patterns of closure duration for stop consonants (Rosen, 1992). However, reversal of speech wave forms in time does not eliminate their having a speech-like quality; many aspects of forward speech are preserved, such as F0, formant transitions, and frication.

While adults do show impaired discrimination of time-reversed stop syllables (Li and Boothroyd, 1996) when perception is measured, the use of time-reversed speech as a masker does not markedly alter performance as compared with a forward speech interferer (Brungart and Simpson, 2002b; Hawley *et al.*, 2004). In contrast, for children who participated in the current study, reversed speech produced different findings than the forward speech. This may be because the reversed speech masker provides more informational masking for children than adults. The general idea that SRM is greater with maskers that have more informational masking has been discussed in recent years by others (Arbogast *et al.*, 2002; Freyman *et al.*, 2001). In those studies the

focus was on comparison of speech and noise-type maskers. Results presented here suggest that informational masking contributes to greater SRM in children, but the difference is borne out of comparisons between reversed-speech and noise type maskers.

E. SRM is independent of female voice spectra

This study examined the effect of two different female voices on masking and SRM in children and adults. The spectra of female voices are notoriously variable, in particular with regard to energy components that are high frequency. Therefore, the question of whether variability in energetic masking affects SRM is highly relevant to studies on the cocktail party effect. We have demonstrated that, although the amount of masking produced by a female voice containing spectral energy approximating that of an “average” female voice (Stelmachowicz *et al.*, 1993) is greater than that for a female voice with reduced spectral energy in the high frequencies, the ability of listeners to utilize spatial cues for source segregation remains the same. This is because the amount of masking produced by Female B leads to increase in the Front and Right SRTs that are virtually identical, for both children and adults. SRM is therefore not dependent directly on the type of female voice used, rendering this measure quite robust under several conditions.

V. CONCLUSIONS

Under conditions studied here, children generally perform worse than adults in the presence of a single interferer. Overall, children have higher SRTs and experience greater masking. In addition, masking and SRM vary with the type of interfering sound, and this variation interacts with age. Compared with adults, children experience greater amounts of masking, and also SRM, especially in the presence of a time-reversed speech interferer. This is quite interesting, because, although this interferer is probably not encountered in everyday listening situations, it may be akin to novel sounds such as foreign languages, which are indeed a part of children’s realistic auditory environments. Finally, the exact voice that is used for measuring these effects contributes to the amount of energetic masking, but does not affect the benefit of spatially segregating target speech from interferers. One of the long-term goals of this work is to identify situations that enable some level of prediction about children’s ability to function in complex environments. Ultimately, such measures can potentially be applied in clinical settings to assess performance of children with hearing impairments and hearing prosthetic devices.

ACKNOWLEDGMENTS

This research was supported by the NIH-NIDCD (Grant Nos. R01-DC003083 and R21-DC05469 to R.Y.L.). The authors would like to thank Allison Olson and Shelly Godar for helping with data collection, Dr. Mary Lindstrom for assistance with statistical analyses, and Dr. Robert Lutfi for suggestions regarding deviance measurements for psychometric functions. We are very grateful to the children and adults who participate in our “listening games.”

¹The deviance measure is $D=2\sum_{i=1}^K\{n_i p_i \log(p_i/p_i) + n_i(1-p_i)\log((1-p_i)/(1-p_i))\}$ where K equals the number of data points of each psychometric function, and n_i is the number of trials for the i th data point.

- Allen, P., and Wightman, F. (1995). “Effects of signal and masker uncertainty on children’s detection,” *J. Speech Hear. Res.* **38**, 503–511.
- ANSI (1987). ANSI S3.9–1987, *American National Standards Specification for Instruments to Measure Aural Acoustic Impedance and Admittance* (American National Standards Institute, New York).
- ANSI (1989). ANSI S3.9–1989, *American National Standards Specification for Audiometers* (American National Standards Institute, New York).
- Arbogast, T. L., Mason, C. R., and Kidd, G. (2002). “The effect of spatial separation on informational and energetic masking of speech,” *J. Acoust. Soc. Am.* **112**, 2086–2098.
- Berman, S., and Friedman, D. (1995). “The development of selective attention as reflected by event-related potentials,” *J. Exp. Child Psychol.* **59**, 1–31.
- Bronkhorst, A. (2000). “The cocktail party phenomenon: A review of research on speech intelligibility in multiple-talker conditions,” *Acoustica* **86**, 117–128.
- Bronkhorst, A. W., and Plomp, R. (1988). “The effect of head-induced interaural time and level differences on speech intelligibility in noise,” *J. Acoust. Soc. Am.* **83**, 1508–1516.
- Bronkhorst, A. W., and Plomp, R. (1992). “Effect of multiple speechlike maskers on binaural speech recognition in normal and impaired hearing,” *J. Acoust. Soc. Am.* **92**, 3132–3139.
- Brungart, D. S., and Simpson, B. D. (2002a). “The effects of spatial separation in distance on the informational and energetic masking of a nearby speech signal,” *J. Acoust. Soc. Am.* **112**, 664–676.
- Brungart, D. S., and Simpson, B. D. (2002b). “Within-ear and across-ear interference in a cocktail-party listening task,” *J. Acoust. Soc. Am.* **112**, 2985–2995.
- Brungart, D. S., Simpson, B. D., Ericson, M. A., and Scott, K. R. (2001). “Informational and energetic masking effects in the perception of multiple simultaneous talkers,” *J. Acoust. Soc. Am.* **110**, 2527–2538.
- Cherry, E. C. (1953). “Some experiments on the recognition of speech, with one and two ears,” *J. Acoust. Soc. Am.* **25**, 975–979.
- Cranford, J. L., Morgan, M., Scudder, R., and Moore, C. (1993). “Tracking of ‘moving’ fused auditory images by children,” *J. Speech Hear. Res.* **36**, 424–430.
- Culling, J. F., Hawley, M. L., and Litovsky, R. Y. (2004). “The role of head-induced interaural time and level differences in the speech reception threshold for multiple interfering sound sources,” *J. Acoust. Soc. Am.* **116**, 1057–1065.
- Dirks, D. D., and Wilson, R. H. (1969). “The effect of spatially separated sound sources on speech intelligibility,” *J. Speech Hear. Res.* **12**, 650–664.
- Doyle, A. (1973). “Listening to distraction: A developmental study of selective attention,” *J. Exp. Child Psychol.* **15**, 100–115.
- Durlach, N. I., Mason, C. R., Shinn-Cunningham, B. G., Arbogast, T. L., Colburn, H. S., and Kidd, G., Jr. (2003). “Informational masking: Counteracting the effects of stimulus uncertainty by decreasing target-masker similarity,” *J. Acoust. Soc. Am.* **114**(1), 368–379.
- Eggermont, J. J., Brown, D. K., Ponton, C. W., and Kimberley, B. P. (1996). “Comparison of distortion product otoacoustic emission (DPOAE) and auditory brain stem response (ABR) traveling wave delay measurements suggests frequency-specific synapse maturation,” *Ear Hear.* **17**, 386–394.
- Feston, J. M., and Plomp, R. (1990). “Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing,” *J. Acoust. Soc. Am.* **88**, 1725–1736.
- Freyman, R. L., Balakrishnan, U., and Helfer, K. S. (2001). “Spatial release from informational masking in speech recognition,” *J. Acoust. Soc. Am.* **109**, 2112–2122.
- Freyman, R. L., Halfer, K. S., McCall, D. D., and Clifton, R. K. (1999). “The role of perceived spatial separation in the unmasking of speech,” *J. Acoust. Soc. Am.* **106**, 3578–3588.
- Gomes, H., Malholm, S., Christodoulou, C., Ritter, W., and Cowan, N. (2000). “The development of auditory attention in children,” *Front. Biosci.* **5**, d108–d120.
- Green, D. M., and Swets, J. (1974). *Signal Detection Theory and Psychophysics* (Kreger, New York).
- Grose, J. H., Hall, J. W., III, and Gibbs, C. (1993). “Temporal analysis in children,” *J. Speech Hear. Res.* **36**, 351–356.
- Guttentag, R. E., and Ornstein, P. A. (1990). “Attentional capacity and chil-

- dren's memory strategy use," in *Development of Attention: Research and Theory*, edited by J. T. Enns (Elsevier, Amsterdam). pp. 305–320.
- Hall, J. W., III, and Grose, J. H. (1994). "Development of temporal resolution in children as measured by the temporal modulation transfer function," *J. Acoust. Soc. Am.* **96**, 150–154.
- Hall, J. W., III, Grose, J. H., Buss, E., and Dev, M. B. (2002). "Spondee recognition in a two-talker masker and a speech-shaped noise masker in adults and children," *Ear Hear.* **23**, 159–165.
- Hartley, D. E., Douglas, E. H., Wright, B. A., Hogan, S. C., and Moore, D. R. (2000). "Age-related improvements in auditory backward and simultaneous masking in 6- to 10-years-old children," *J. Speech Lang. Hear. Res.* **43**, 1402–1415.
- Hawley, M. L., Litovsky, R. Y., and Colburn, H. S. (1999). "Speech intelligibility and localization in a multi-source environment," *J. Acoust. Soc. Am.* **105**, 3436–3448.
- Hawley, M. L., Litovsky, R. Y., and Culling, J. F. (2004). "The benefit of binaural hearing in a cocktail party: Effect of location and type of interferer," *J. Acoust. Soc. Am.* **105**, 3436–3448.
- Hazan, V., and Markham, D. (2004). "Acoustic-phonetic correlates of talker intelligibility for adults and children" *J. Acoust. Soc. Am.* **116**, 3108–3118.
- Kidd, G., Mason, C. R., and Arbogast, T. L. (2002). "Similarity, uncertainty, and masking in the identification of nonspeech auditory patterns," *J. Acoust. Soc. Am.* **111**, 1367–1376.
- Leek, M. R., Brown, M. E., and Doorman, M. F. (1991a). "Informational masking and auditory attention," *Percept. Psychophys.* **50**, 205–214.
- Leek, M. R., Thomas, E. H., and Marshall, L. (1991b). "An interleaved tracking procedure to monitor unstable psychometric functions," *J. Acoust. Soc. Am.* **90**, 1385–1397.
- Levitt, H. (1971). "Transformed up-down methods in psychophysics," *J. Acoust. Soc. Am.* **49**, 467–477.
- Li, A. C., and Boothroyd, A. (1996). "Speech perception of temporally reversed syllables by normally hearing adults" (in Chinese), *Zhonghua Yi Xue Za Zhi, (Taipei)* **57**, 1–6.
- Litovsky, R. Y. (2005). "Speech intelligibility and spatial release from masking in young children," *J. Acoust. Soc. Am.* **117**, 3091–3099.
- Litovsky, R. Y. (2004). "Method and system for rapid and reliable testing of speech intelligibility in children," *J. Acoust. Soc. Am.* **115**, 2699.
- Litovsky, R. Y. (1997). "Developmental changes in the precedence effect: Estimates of minimum audible angle," *J. Acoust. Soc. Am.* **102**, 1739–1745.
- Litovsky, R. Y., and Ashmead, D. (1997). "Developmental aspects of binaural and spatial hearing," in *Binaural and Spatial Hearing*, edited by R. H. Gilkey and T. R. Anderson (Earlbaum, Hillsdale, NJ) pp. 571–592.
- Lutfi, R. A., (1990). "How much masking is informational masking?" *J. Acoust. Soc. Am.* **88**, 2607–2610.
- Lutfi, R. A., Kistler, D. J., Callahan, M. R., and Wightman, F. L. (2003). "Psychometric functions for informational masking," *J. Acoust. Soc. Am.* **114**, 3273–3282.
- Macoby, E. (1969). "The development of stimulus selection," *Minn. Symp. Child Psych.* **3**, 68–96.
- Morrongiello, B. A., Kulig, J. W., and Clifton, R. K. (1984). "Developmental changes in auditory perception," *Child Dev.* **55**, 461–71.
- Papso, C. F., and Blood, I. M. (1989). "Word recognition skills of children and adults in background noise," *Ear Hear.* **10**, 235–236.
- Pessig, J., and Kollmeier, B. (1997). "Directivity of binaural noise reduction in spatial multiple noise-source arrangements for normal and impaired listeners," *J. Acoust. Soc. Am.* **101**, 1660–1670.
- Plomp, R., and Mimpen, A. M. (1981). "Effect of the orientation of the speaker's head and the azimuth of a noise source on the speech reception threshold for sentences," *Acustica* **48**, 325–328.
- Plude, D. J., Enns, J. T., and Brodeur, D. (1994). "The development of selective attention: A life-span overview," *Acta Psychol.* **86**, 227–272.
- Ponton, E. W., Eggermont, J. J., Coupland, S. G., and Winkelaar, R. (1992). "Frequency-specific maturation of the eighth nerve and brain-stem auditory pathway: Evidence from derived auditory brain-stem responses (ABRs)," *J. Acoust. Soc. Am.* **91**, 1576–1586.
- Ramos, F., Hauser, M. D., Miller, C., Morris, D., and Mehler, J. (2000). "Language discrimination by human newborns and by cotton-top tamarin monkeys," *Science* **288**, 349–352.
- Rosen, S. (1992). "Temporal information in speech: Acoustic, auditory and linguistic aspects," *Philos. Trans. R. Soc. London, Ser. B* **336**, 367–373.
- Rothauer, E. H., Chapman, W. D., Guttman, N., Nordby, K. S., Silbigert, H. R., Urbanek, G. E., and Weinstock, M. (1969). "IEEE Recommended practice for speech quality measurements," *IEEE Trans. Audio Electroacoust.* **17**, 225–246.
- Stelmachowicz, P., Mace, A. L., Kopun, J. G., and Carney, E., (1993). "Long-term and short-term characteristics of speech: Implications for hearing aid selection for young children," *J. Speech Hear. Res.* **36**, 609–620.
- Veloso, K., Hall, J. W., III, and Grose, J. H. (1990). "Frequency selectivity and comodulation masking release in adults and in 6-years-old children," *J. Speech Hear. Res.* **33**, 96–102.
- von Békésy, G. (1960). *Experiments in Hearing* (McGraw-Hill, New York).
- Watson, C. S., Kelly, W. J., and Wroton, H. W. (1976). "Factors in the discrimination of tonal patterns. II. Selective attention and learning under various levels of stimulus uncertainty," *J. Acoust. Soc. Am.* **60**, 1176–1186.
- Wichmann, F. A., and Hill, J. (2001a). "The psychometric function. I. Fitting, sampling, and goodness of fit," *Percept. Psychophys.* **63**, 1290–1313.
- Wichmann, F. A., and Hill, J., (2001b). "The psychometric function. II. Bootstrap-based confidence intervals and sampling," *Percept. Psychophys.* **63**, 1314–1329.