Published in final edited form as:
J Acoust Soc Am. 2005 November ; 118(5): 3164.

Informational masking of speech in children: Effects of ipsilateral and contralateral distracters

Frederic L. Wightman[a] and **Doris J. Kistler**
Heuser Hearing Institute and Department of Psychological and Brain Sciences, University of Louisville, Louisville, Kentucky 40292

Abstract

Using a closed-set speech recognition paradigm thought to be heavily influenced by informational masking, auditory selective attention was measured in 38 children (ages 4–16 years) and 8 adults (ages 20–30 years). The task required attention to a monaural target speech message that was presented with a time-synchronized distracter message in the same ear. In some conditions a second distracter message or a speech-shaped noise was presented to the other ear. Compared to adults, children required higher target/distracter ratios to reach comparable performance levels, reflecting more informational masking in these listeners. Informational masking in most conditions was confirmed by the fact that a large proportion of the errors made by the listeners were contained in the distracter message(s). There was a monotonic age effect, such that even the children in the oldest age group (13.6–16 years) demonstrated poorer performance than adults. For both children and adults, presentation of an additional distracter in the contralateral ear significantly reduced performance, even when the distracter messages were produced by a talker of different sex than the target talker. The results are consistent with earlier reports from pure-tone masking studies that informational masking effects are much larger in children than in adults.

I. INTRODUCTION

The ability to attend selectively to individual objects in the auditory world is undeniably one of the most important and complex skills possessed by hearing adults. In spite of the fact that relevant and irrelevant auditory events are commingled acoustically before arriving at the ears, the auditory system and the brain can, in most instances, parse the auditory “scene.” This auditory source segregation process requires identification and segregation of the constituents of a scene, attention to one target sound source, and suppression of other distracting sources. For example, at a social gathering, successful interaction requires that we listen to the individual with whom we are talking while simultaneously disregarding the clamor of other voices around us. Similarly, a child must be able to attend to his or her teacher while disregarding the various sources of noise in the classroom. The mechanisms and processes that subserve source segregation and selective attention in adults are not yet fully understood, in spite of a substantial research effort over the last quarter century. [See Bregman (1990) for a comprehensive introduction to auditory source segregation.] Relatively little research has focused on the development of either auditory source segregation or auditory selective attention in children.

Experiments that measure detection or discrimination of tones in the presence of simple maskers or distracters (other tones or noise) tap a very basic form of auditory source segregation and selective attention. Even in these simple tasks, children up to the ages of 6 to 7 years perform more poorly than adults (Allen, 1991; Allen and Nelles, 1996; Allen and Wightman,

---

[a] fred.wightman@louisville.edu.

1992, 1995; Allen *et al.*, 1989; Bargones and Werner, 1994; Bargones, Werner, and Marean, 1995; Schneider *et al.*, 1989; Stellmack *et al.*, 1997; Wightman and Allen, 1992; Willihnganz *et al.*, 1997). However, simple tone detection paradigms do not present the kind of source segregation challenge that we face in real life, primarily because the "distraction" is typically relatively static and has a very different quality than the target signal. It is now well documented that, if a detection paradigm is modified to introduce uncertainty and signal-like quality in the distracter stimulus, the segregation/attention problem becomes much more difficult. For example, when adult listeners are asked to detect a tone masked by a complex of other tones, randomizing the frequencies and levels of the masker tones produces a significant decrement in performance, represented by as much as a 40 dB threshold elevation in some conditions and some listeners (Neff and Callaghan, 1988; Oh and Lutfi, 1998). This effect has been called "informational masking" to distinguish it from "energetic masking," which is masking thought to be caused by overlap of signal and masker spectra. Large individual differences characterize informational masking data. Some adult listeners show little or no informational masking, while others show a great deal (Oh and Lutfi, 1998). Recent results from our laboratory (Oh, Wightman, and Lutfi, 2001; Wightman *et al.*, 2003) suggest that informational masking is much greater in young children. Moreover, the individual differences in the data from children of the same age are also large. Even among teenagers, the data from some are like preschoolers and from others are like adults (Wightman *et al.*, 2003).

The recent models of informational masking, developed in the context of the pure-tone detection paradigms, suggest that listeners base their detection decisions on the weighted sum of outputs of a bank of auditory filters or channels (Lutfi, 1989, 1993; Lutfi *et al.*, 2003a; Oh and Lutfi, 1998; Richards, Tang, and Kidd, 2002; Tang and Richards, 2003; Wright and Saberi, 1999). The large differences in informational masking between children and adults and the large individual differences in thresholds are well predicted by the version of the model proposed by Oh and Lutfi (1998). In this model individual differences are explained by changes in channel weights and in "weighting efficiency," a measure that represents the difference between a listener's weights and those of an ideal observer in the same task. In most cases, the ideal observer would place nearly all weight on the signal channel and little or no weight on the channels passing the masker components. Data from our laboratory suggest that children display more informational masking because they place much more weight than adults on the nonsignal channels (Oh *et al.*, 2001). In other words, children appear to listen less selectively than adults.

The hypothesis that children listen less selectively than adults is generally consistent with the data from numerous studies of the development of selective attention as assessed in dichotic listening paradigms. In a typical dichotic listening experiment, a listener is asked to attend to the message presented to one ear and to ignore a distracting message presented in the other ear. Two kinds of results emerge from these studies. First, there is a consistent "right-ear advantage" (REA) which is revealed by higher performance when the target is presented to the right ear than when it is in the left ear. Most of the evidence suggests that the magnitude of the REA is the same in children as in adults (Bryden and Allard, 1981; Geffen, 1976, 1978; Geffen and Sexton, 1978; Hiscock and Kinsbourne, 1977, 1980; Sexton and Geffen, 1979), but this is still a controversial issue (Hugdahl, Carlsson, and Eichele, 2001; Morris *et al.*, 1984; Pohl, Grubmuller, and Grubmuller, 1984). A second consistent finding is that children do not perform as well as adults in dichotic listening tasks. The poorer performance is a result of more intrusions from the distracter message in the opposite ear, thus suggesting poorer attentional selection in children. This is taken as evidence that some of the brain processes that mediate performance in an auditory selective attention task are not mature in children (Doyle, 1973; Maccoby, 1969; Maccoby and Konrad, 1966). Data from recent studies of event-related brain potentials evoked during a selective attention task (Berman and Friedman, 1995; Coch, Sanders, and Neville, 2005) are consistent with this view. These results suggest that the N1 attention effect first reported by Hillyard *et al.* (1973) is not adult-like until after age 8 (Coch *et al.*, 2005). However, it is not clear whether this developmental effect originates in mechanisms of executive control, sustained attention, stimulus selection, or some other process that subserves selective attention (Gomes *et al.*, 2000).

A previous study of dichotic listening using pure tones as targets and distracters (Wightman *et al.*, 2003) reports results that are qualitatively consistent with those from the above-cited studies. When the target is presented to one ear and the distracter to the other, children produce much higher detection thresholds (poorer performance) than adults. These results, as well as the results from studies in which a pure-tone target and distracter are presented to the same ear (Lutfi *et al.*, 2003a, 2003b; Oh and Lutfi, 1998; Oh *et al.*, 2001), offer a useful picture of the developmental course of pure-tone informational masking. However, it is not clear how the results from pure-tone detection experiments might generalize to a more realistic everyday selective attention task.

Attending to a target speech signal in the presence of other speech distracters is a good example of an everyday auditory selective attention task. Listening to speech is probably the primary vehicle for learning among young hearing children, and speech is rarely heard in a quiet environment. Given the pure-tone detection results one might predict that children would have more difficulty than adults with this task. Although there is some support for this prediction in the literature, the details have yet to be revealed. The dichotic listening literature (cited earlier) offers some clues, but only for the unnatural case in which target and distracter are in different ears. Data from two recent studies (Fallon, Trehub, and Schneider, 2000; Hall *et al.*, 2002) do suggest that children have more difficulty than adults attending to speech that is presented with a speech distracter in the same ear.

One complexity of the speech task is the fact that both energetic and informational masking are likely to be involved. Because interfering speech overlaps the target both temporally and spectrally, energetic masking is to be expected. However, the focus here is on informational masking, which in the speech task would be caused both by the similarity of the target and distracter speech signals and by the uncertainty of the distracter from trial to trial. It is difficult, if not impossible, to disentangle energetic and informational masking effects completely, but previous research offers some useful techniques that may be applied to studies with children.

A series of experiments recently reported by Brungart and colleagues (Brungart, 2001a, 2001b; Brungart and Simpson, 2002, 2004; Brungart *et al.*, 2001; Kidd *et al.*, 2003) uses a novel speech intelligibility task that offers sensitive measures of “purely” informational masking effects. The “Coordinate Response Measure” (CRM) is a task in which listeners are asked to attend to a spoken target message of the form, “Ready *call sign*, go to *color number* now.” The target “call sign” is fixed for a block of trials (or for the entire experiment) and is chosen from a group of eight (e.g., “baron,” “ringo,” “arrow,” “tiger”). The target “color” and “number” are chosen randomly for each trial, from a set of four colors (red, blue, green, and white) and eight numbers (1–8). One or more distracter messages is mixed with the target message and has exactly the same form as the target but different call sign(s), color(s), and number(s). Target and distracter messages are temporally aligned at the beginning, and are roughly the same total duration. Typically the target talker (one of eight, four males and four females) is fixed for a block of trials (often the entire experiment), and the distracter talker(s) is either the same as or different from the target. If different, the distracter can be either the same or the opposite sex. In the CRM task listeners are asked to report the target color and number on a computer screen. The ratio (T/D, in dB) of overall target rms to overall ipsilateral distracter rms is varied, and percent correct is recorded at each T/D. The articles describing the development and early experiments with the CRM (Bolia *et al.*, 2000; Brungart, 2001a) describe the task and stimulus materials in detail and show how the results compare with other more traditional measures of speech intelligibility.

The experiments described by Brungart (2001b) provide convincing evidence that masking in the CRM task is dominated by informational masking. The initial experiments investigated the masking effects of a single distracter presented in the same ear as the target. When the distracter was a speech-shaped noise with its envelope modulated by a CRM message, performance was considerably better (near perfect performance down to T/D of −6 dB) than with any of the speech distracters. The noise presumably approximated the energetic masking components of the speech distracter stimuli. Since performance was much poorer with the speech distracter, the assumption is that the overall masking effect with the speech distracter was primarily informational. When the distracter and target speakers were of opposite sexes, there was a dramatic improvement in performance over the condition in which both target and distracter speakers were of the same sex. This was most likely a result of the fact that male and female voices were perceptually very different, thus reducing informational masking. There was a smaller performance improvement when the target and distracter talkers were different, but of the same sex. The same arguments could be made in this case about reductions in target and distracter similarity. Finally, indications of the dominance of informational masking in the CRM task were found in the shapes of the psychometric functions (% correct vs T/D) in the various conditions. With a modulated noise distracter the psychometric functions were monotonic ogives, from chance (3% correct) to perfect performance, typical of detection or other speech intelligibility results. However, when the distracter was speech, the psychometric function displayed a prominent plateau from about 0 dB T/D to about −10 dB T/D. In other words, over a 10 dB range in which the target was progressively less intense than the distracter, performance remained constant. Brungart explained this in the context of informational masking. Especially in the conditions in which the target and distracter talkers were the same, intensity (loudness) could have provided information that might have allowed the listener to segregate target and distracter, since all other attributes of the two voices were the same. Thus, when the target was less intense than the distracter, so long as it was audible, a listener might be able to adopt the strategy of “listening to the softer voice.” This strategy is clearly inconsistent with the principles of energetic masking.

The assertion that the masking effects of a speech distracter in the CRM task are primarily informational is strengthened by a detailed analysis of the listeners’ errors. Brungart (2001b) shows that the color or number errors are not random, as might be expected if the masking were primarily energetic. Rather, the vast majority of errors come from the distracter message. For example, in one condition (same sex distracter) approximately 70% of the number responses were correct at a T/D of 0 dB. From the 30% number errors at this T/D ratio, approximately 28% (thus, nearly all of them) were from the distracter message. This suggests that listeners could hear both target and distracter messages but made errors because of an inability either to segregate the two messages or to remember which message was the target.

The experiment reported here explores informational masking in children using the CRM task. The aim is to understand both the extent and character of informational masking with speech stimuli and the development changes in the effects that occur from preschool age to young adulthood.

II. METHODS

A. Listeners

Eight adults and 38 children served as participants in this experiment. An additional 3 adults were recruited but did not complete the experiment. Preschool children were recruited from the University of Wisconsin’s Waisman Early Childhood Program. Older children and adults were recruited from the University of Wisconsin community. The adults ranged in age from 20 to 30 years. For convenience in data interpretation, the children were divided into five age groups: 8 children were in the age range 4.6–5.7 years; 6 in the 6.6–8.5 year group; 8 in the 9.6–11.5 year group; 8 in the 11.6–13.5 year group; 8 in the 13.6–16.0 group. All adults and children passed a 20 dB HL screening for hearing loss. The preschool (4.6–5.7 years) children were tested for middle-ear problems (routine tympanometry) before each session, and the other children were tested similarly before the first session and irregularly thereafter. Sessions were canceled whenever there were indications of middle-ear problems. All children who were recruited completed the experiment.

B. Stimuli

The target and distracter speech messages were drawn from the corpus of CRM stimuli made available by Bolia *et al.* (2000). The corpus includes 2048 phrases of the form, “Ready, *call sign*, go to *color number* now.” Eight talkers (4 male, 4 female) are recorded, each speaking 256 different phrases (eight call signs, “baron,” “ringo,” “tango,” etc.; eight numbers, 1–8; four colors, “red,” “white,” “green,” “blue”). For this experiment, the target phrase was always spoken by talker 1 (male) from the corpus, using the call sign “baron.” The distracter phrases always used a different talker, call sign, color, and number. In some conditions the sex of the distracter talker was male, and in other conditions female. With the exception of the distracter talker’s sex, the distracter phrases(s) used on each trial were chosen randomly from the available, nontarget phrases.

All conditions involved trials in which a single target and single distracter were presented to the listener’s right ear. In some conditions an additional distracter phrase was presented to the left ear. Another condition involved presentation of a Gaussian noise with a speech-shaped spectrum to the left ear. The overall level (rms) of the distracter in the target ear was held constant for all conditions at approximately 62 dB SPL. When a speech distracter was present in the nontarget ear, it was presented at this same rms level. When a noise distracter was present in the nontarget ear, its rms level was 77 dB SPL. The rms level of the target was varied from trial to trial. Depending on listener and condition the target/distracter ratio (T/D) ranged from +24 to −24 dB. Thus the highest rms level of the target was 86 dB SPL.

The stimulus materials were produced digitally (CRM stimuli taken from the distribution CD, noise synthesized on a PC), converted to analog form (44.1 kHz sample rate) by Tucker-Davis Technologies (TDT) DD1 D/A converters, mixed, amplified and presented to listeners via calibrated Beyer DT990-Pro headphones. Target and distracter levels were controlled by programmable attenuators (TDT PA-4) prior to mixing. The target and distracter phrases were time-aligned on the distribution CD such that the word “ready” for target and distracter phrases started at the same time.

C. Conditions

The experiment involved three distracter conditions. In all conditions a target phrase and a simultaneous distracter phrase (different talker, call sign, color, and number than the target) were presented to the listener’s right ear. A monaural condition involved no stimulus presented to the contralateral ear. The second distracter condition added a speech-shaped noise distracter to the contralateral ear. In the third distracter condition a speech distracter (CRM phrase with different talker, call sign, color, and number than either the target or the ipsilateral distracter) was presented to the contralateral ear. Each distracter condition was tested with both male and female distracters. Thus, there were six total conditions in the experiment. The youngest children were not tested with male distracters since pilot testing suggested that, with a contralateral distracter stimulus, target levels in excess of 90 dB SPL would be required to obtain above-chance performance.

D. Procedure

Listeners sat in a sound-isolated room (IAC 1200) in front of a computer display. The display showed a start button and 32 response buttons arranged in four colored matrices of 8 buttons each, numbered 1–8. Individual trials were initiated by the listener by a mouse-click on the start button, except for sessions involving the youngest (preschool) children, in which case the experimenter (seated beside the child) initiated the trial once the listener was quiet and appeared attentive. After hearing the phrases, the listener moved the mouse cursor to the matrix of the heard color and clicked on the number corresponding to the heard number. The preschoolers pointed to the color-number of the heard phrase and the experimenter entered the response. Visual feedback was given (“smiley face” on computer screen) for a correct response. The preschool children were also verbally reinforced after each trial.

The level of the target was varied in an up-down staircase fashion (5 levels) from trial to trial so that an entire psychometric function, from near perfect performance to chance, could be estimated during each session. The highest level was presented first. Then the level was decreased on each of the next 4 trials, after which it was increased for 4 trials. This down-then-up sequence was then repeated for the duration of the run. In this way the highest and lowest levels were tested on half as many trials as the intermediate levels. For the preschool children, the 5 levels were 4 dB apart, so a 16 dB range was covered. For the older children and adults, the levels were 8 dB apart, and the total range of target levels was 32 dB. The highest target level was chosen (based on pilot testing with each listener) in order to assure near perfect performance. Practice runs with no distracter were also tested to assure perfect performance at each target level. In any given run only 5 levels were presented, but for some adults and children the starting levels were different on different runs to evaluate both perfect and near-chance performance.

Both children and adults completed at least 1 practice run in each of the six conditions (and one with no distracter) before data were collected. Children were tested in an additional practice run if it appeared necessary. Most young children were tested in two practice runs. Following practice, runs from the six conditions were tested in a pseudorandom order, such that for each run a random choice was made among all the remaining runs across all conditions.

For children the total number of trials in each condition ranged from 256 to 384. For adults, the total number of trials in each condition ranged from 512 to 640. Testing was conducted over the course of several sessions. The preschool children were tested in 32-trial runs, and they would normally complete 2 such runs in a session. Older children and adults were tested in 64 or 128 trial runs depending on the listener. For the preschool children, the sessions were always less than 30 min and were held once per week. Completion of the experiment required as many as 12 sessions for these children. The adult sessions were 1.5–2.0 h long and were scheduled at the participant’s convenience, usually twice per week. Most adults completed the experiment in 4–6 sessions. For children at intermediate ages, sessions were 1–2 h, depending on the participant, and the number of sessions required to complete the experiment ranged from 3 to 5. Frequent breaks were encouraged for all participants.

Children in the youngest age group were given a small toy at the end of each session. All other participants were paid $7/h for their participation.

III. RESULTS AND DISCUSSION

A. Data analysis

Interpreting psychometric function data from children is complicated by several factors. First, young children cannot be tested as extensively as older children and adults, so comparing results across age groups is not always straightforward. Second, individual differences are often larger in the younger age groups (Oh *et al.*, 2001; Wightman and Allen, 1992; Wightman *et al.*, 2003). Third, the upper asymptote of many psychometric functions obtained from children does not reach 100% correct, even for very strong signals, implying a certain degree of inattention. Finally, in the current paradigm, the same numbers of trials were not run at all T/D ratios because of the nature of the staircase procedure. In previous studies (Wightman *et al.*, 2003) these problems were mitigated by fitting a smooth curve (logistic) to the psychometric function data from each listener in each condition and extracting estimates (with confidence limits) of the parameters of those fitted functions (Wichmann and Hill, 2001a, 2001b). This allowed interpretation of the thresholds (T/D ratio producing 75% correct) for each listener. In the current study, complete psychometric functions were also obtained from each listener in each condition. However, it is not meaningful to compute the parameters of best fitting functions from these data. Many of the functions are irregular in shape, either nonmonotonic or having a plateau, similar to what has been reported before (Brungart *et al.*, 2001; Brungart and Simpson, 2002). For these reasons, we will show complete psychometric functions when discussing the results.

In order to reduce the data set from the 38 children and 8 adults to a manageable size, the children were assigned to age groups spanning roughly 2 years, as described earlier. The psychometric functions from all the children in each of the six age groups and each of the six conditions were plotted (a total of 36 plots) and examined visually for homogeneity. Figure 1 shows four representative examples. The data from individual listeners are plotted along with the means (weighted according to the number of trials included for each listener at each T/D), for two conditions and two age groups. Our conclusion from this informal analysis was that in spite of large individual differences in percent correct at specific T/D ratios, the shapes of the individual psychometric functions in each age group are well represented by the mean. Thus, all further analyses are based on the group mean psychometric function in each condition.

Although the experimental design did not allow a systematic assessment of potential training effects, an informal analysis suggested that, even for the youngest children tested, performance was approximately constant for the duration of the experiment.

B. Male distracter conditions

Figure 2 shows mean psychometric functions obtained from five of the six age groups (the preschool children were not tested in these conditions) in the three male distracter conditions. To facilitate comparison across age groups, only partial psychometric functions are shown. It is important to note in this context that all listeners reached near 100% correct performance at high T/D ratios and approached chance at low T/D ratios. The stimulus configurations represented in this figure are nearly identical to those in the Brungart and Simpson (2002) study of adults. The main differences are that the Brungart and Simpson experiment used a contralateral noise distracter 5 dB more intense than what was presented here. Consider first the current data from the adults. These are entirely consistent with those shown in Fig. 2 of Brungart and Simpson (2002). With no distracter in the contralateral ear, performance is nearly perfect at a T/D ratio of +8 dB, decreases to about 70% correct at T/D of 0 dB, remains at 70% until a T/D of about −8 dB, and falls for lower T/D ratios. The plateau in performance between a T/D of 0 and −8 dB has been reported before (Dirks and Bower, 1969;Egan, Carterette, and Thwing, 1954) and probably reflects these listeners’ abilities to segregate target and distracter on the basis of intensity. The addition of a noise distracter to the contralateral ear has no effect, also as shown in the Brungart and Simpson (2002) experiment, but the addition of a contralateral speech distracter has a substantial negative effect, reducing performance by as much as 30% at some T/D ratios. Even though, as shown by Brungart and Simpson (2002), a contralateral speech distracter has no impact when presented alone, its simultaneous presence with an ipsilateral distracter apparently interferes with listeners' abilities to segregate the target and the ipsilateral distracter.

The mean data from the four groups of children (Fig. 2) suggest increasing amounts of informational masking in all conditions as age decreases. At a T/D ratio of +8 dB, performance is quite good (over 80%) in all age groups, but the youngest group (6.6–8.5 years) obtains less than 35% correct at the 0 dB T/D ratio. The performance plateau between T/D ratios of 0 and −8 dB is not evident in the data from the youngest group but is obvious in the data from the other groups. Thus, from these data, we might conclude that only listeners 9 years of age or older use the intensity-based target-distracter segregation strategy. The lack of an effect caused by the additive contralateral noise appears in all age groups. Finally, the substantial decrement in performance caused by the addition of a contralateral speech distracter is roughly constant across the age groups, but slightly smaller for the youngest group tested (probably a floor effect). The constancy of the contralateral speech effect is somewhat surprising. Even though the data from this and previous experiments (Oh *et al.*, 2001;Wightman *et al.*, 2003) suggest that informational masking decreases as the age of the listener increases, the added component of informational masking contributed by a contralateral distracter seems not to be age dependent.

The age effects in each of the three male distracter conditions are shown more clearly in Fig. 3. This figure shows the same data as Fig. 2, but displayed differently. The three panels represent the three distracter conditions and the different symbols represent the different age groups. Note that with few exceptions, at each T/D ratio performance decreases monotonically with decreasing age, although there is virtually no difference in performance between the two oldest age groups of children. Note also that the plateau effect is evident in both the monaural and contralateral noise conditions in all age groups except the youngest tested.

The masking effects shown in Figs. 2 and 3 are almost certainly dominated by informational, not energetic, masking. Brungart (2001b) offers several kinds of evidence supporting this assertion. First, when the ipsilateral distracter in his experiment was a speech-shaped noise modulated by the envelope of a CRM phrase, adult listeners' performance was near perfect at a T/D ratio of 0 dB, only decreasing to about 60% at a T/D of −8 dB. It is reasonable to assume that the masking produced by the modulated noise would represent an upper bound on energetic masking in these conditions. Thus, the masking effects observed in the Brungart (2001b) experiment, at least down to T/D of −8 dB, would be primarily informational. Brungart (2001b) also presented an analysis of errors that strongly implicated informational masking. The overwhelming majority of color and/or number errors made by his adult listeners with an ipsilateral speech distracter consisted of colors and/or numbers that were present in the distracter phrases. If energetic masking were at work, one might expect the listener to choose a response randomly, thus producing a chance distribution of errors. A similar analysis of errors for the adult listeners in the current experiment is shown in Fig. 4. Color and number errors are combined, and the number of color and/or number errors expected to be present by chance (random errors) in the distracter is shown by the dashed lines. Note that in both the monaural and contralateral noise conditions (upper panels), for T/D ratios greater than −8 dB, more than 90% of the total responses (both correct and incorrect) were contained in either the target or distracter phrases. The same is true for the contralateral speech condition (lower panels). The data from the contralateral speech condition are plotted twice in this figure to clarify the number of errors expected by chance to be contained in both the ipsilateral and contralateral distracters. The contralateral plot (right panel) shows that for T/D ratios between −8 and +8 dB most of the errors came from the ipsilateral distracter, as represented by the fact that the number of errors from the contralateral distracter at these T/D ratios is less than would be expected by chance. The ipsilateral plot (left panel) confirms this by showing that the number of errors contained in the ipsilateral distracter is greater than would be expected by chance. In summary, the fact that at T/D ratios greater than −8 dB most of the errors were contained in the distracter (s) is entirely consistent with what was reported by Brungart (2001b) and argues that the masking effects are primarily informational at these T/D ratios.

The error analyses shown in Figs. 5 and 6 provide evidence that the masking effects for the children are also dominated by informational masking. Figure 5 shows the data from the oldest age group (13.6–16 years). Note that except for a slightly lower performance level (as represented by fewer correct responses) the distribution of responses is indistinguishable from that seen in the adult data (Fig. 4). As was the case in the adult data, for T/D ratios of −8 dB and higher, most of the errors made by the children come from the distracter(s). In addition, most of the errors in the contralateral speech condition come form the ipsilateral distracter. Thus it appears that both the adults and the older children can effectively ignore the distracter in the nontarget ear, since the number of errors contained in the contralateral distracter never rises above that expected by chance. However, the pattern of errors also suggests that both adults and older children can hear and understand both target and distracter messages at T/D ratios of greater than −8 dB, and that the masking effects seen at these T/D ratios are thus informational.

Figure 6 shows the error analysis for the youngest group of children tested in the male distracter conditions. Note that this figure shows no data for T/D ratios below −8 dB; performance for nearly all of these listeners was at chance below this value. Although the pattern of errors is different for these young children than for the adults (Fig. 4) and older children (Fig. 5), the conclusions to be drawn from the error analyses are not fundamentally different. When performance is substantially above chance (at T/D ratios of 0 dB and above), most of the errors are contained in the distracter(s). In the monaural and contralateral noise conditions (upper panels of Fig. 6) the number of errors contained in the distracter is well above chance at T/D ratios of 0 dB and below. In the contralateral speech condition (lower panels of Fig. 6), although performance was near chance at −8 dB T/D ratio, the vast majority of errors at this T/D ratio and above came from the distracter(s). As with the adults and older children most of the errors came from the ipsilateral distracter. As shown in Fig. 6 (lower panels), at T/D ratios above 0 dB the number of errors contained in the contralateral distracter was no greater than would be expected by chance, and at lower T/D ratios was much less than chance. Thus, even the youngest children appear to focus attention on the target ear and thus limit intrusions from the message in the nontarget ear. This result may appear to contradict the findings of our previous study of pure-tone informational masking in children (Wightman *et al.*, 2003), which suggested that children could not ignore the stimulus in the contralateral ear. However, there are many differences between the current experiment and the previous one, not the least of which is the fact that in the current experiment distracters were present in both ears. Thus, the apparent contradiction is not viewed as serious. Moreover, as will be shown in the results from the female distracter conditions, there is evidence here that in some cases children do attend to the contralateral ear when it is not appropriate to do so.

Error analyses for the two intermediate age groups of children are not shown here since the patterns of errors are not substantially different than those shown for the other groups. Performance in all conditions declines systematically with age, as can be seen in Fig. 3, and the error analyses reveal nothing inconsistent with the other error analyses. In all conditions most of the errors for T/D ratios greater than −8 dB are contained in the distracter presented to the target ear.

The main result from the male distracter conditions, that informational masking is greater in children than in adults, agrees with the results of other developmental experiments in the literature on speech recognition with same-sex speech distracters. For example, Hall *et al.* (2002) reported higher speech recognition thresholds in both adults and children with a 2-talker speech distracter than with a noise distracter. This is clear evidence of informational masking. The same authors also reported that the effect is greater in their 5–10 year olds than in their adults. Fallon *et al.* (2000) obtained higher speech recognition thresholds from children than from adults with a multitalker babble distracter, but no control condition was run with a noise distracter to assess the extent of informational masking in this experiment.

C. Female distracter conditions

In the female distracter conditions the ipsilateral and, if present, the contralateral speech distracters were spoken by a female, randomly selected on each trial from the four possible female talkers in the corpus. If there were two distracters they were spoken by two different female talkers. The target talker was the same male (Talker 1 from the corpus) in all conditions. Thus, this condition was similar to the "different sex" condition in the experiment reported by Brungart (2001b), except that in the Brungart study, the sex of the target talker was also allowed to vary from trial to trial.

The mean data from all six age groups (including the preschoolers, who were not tested with male distracters) are shown in Fig. 7. The format of this figure parallels that of Fig. 2. Several features of these data are noteworthy. First, comparing the data from Figs. 2 and 7, it can be seen that for the adults in the monaural condition there is a substantial release from informational masking when the target is male and the distracter is female. At a T/D ratio of 0 dB for example, performance is near perfect with the female distracter and only 70% correct with a male distracter (Fig. 2). Moreover, at a T/D ratio of −16 dB, adult performance with the female distracter is about 75% correct and with a male distracter it was no more than 55%. At this low T/D ratio energetic masking is almost certainly playing a significant role, so a complete release from masking is not to be expected. For the children there is also a substantial release from informational masking when the target talker is male and the distracter talker is female (compare Figs. 2 and 7). At 0 dB T/D ratio with the male distracter, children in the 6.6–8.5 year old group obtained about 30% correct, but with the female distracter performance was above 70% correct. This corresponds to approximately 9 dB of improvement in the T/D ratio required for 50% correct. The release from informational masking in the female distracter conditions was similar in the other groups of children.

Consistent with the diminished contribution of informational masking at low T/D, there is no evidence of the plateau effect that was observed with male distracters at T/D ratios between 0 and −8 dB. The lack of a performance plateau is similar to what was reported by Brungart (2001b) in the "different sex" condition of his experiment. Those data show a gradual decline in performance from 100% at a T/D ratio of +12 dB to about 80% at a T/D of −12 dB, similar to what is shown here for the adult listeners (Fig. 7). It is reasonable to suggest that for the adults and older children, target-distracter segregation is facilitated so much by the difference in target and distracter sex that no further improvement is to be had by the use of the level-segregation strategy.

The impact of the contralateral noise and the contralateral speech distracters in the female distracter conditions was very similar to what was observed with male distracters. Adding noise to the contralateral ear had little or no effect, except at the lowest T/D ratio (−16 dB) where it caused a slight drop in performance in each age group. However, in spite of the release from informational masking in the target ear caused by the use of a different sex distracter, adding female speech to the contralateral ear caused a substantial drop in performance, about 20%, for all groups except the youngest, at T/D ratios near the midpoint of the psychometric functions. The 20% drop was somewhat smaller in some age groups than with male distracters, as might be expected. However, in the 6.6–8.5 year old group the impact of the contralateral female distracter (Fig. 7) appears to be some-what greater than the impact of the contralateral male distracter (Fig. 2). Overall, the effect of adding a contralateral female masker was largely age independent. The youngest children (preschoolers), who were the poorest performers overall, showed a slightly larger effect of the added contralateral female speech masker (30%–40% near the midpoint of the psychometric function).

The age effect in the female distracter conditions is shown in Fig. 8. As in the male distracter conditions (Fig. 3) performance in all conditions declines systematically and monotonically as age decreases. At a T/D ratio of 0 dB for example, the adults scored around 95% correct in the monaural condition, while the children in the youngest age group scored no higher than 30% in the same condition. As in the male distracter conditions, there is almost no difference in performance between the two oldest age groups of children.

An error analysis of the adult data from the female distracter conditions is shown in Fig. 9. These data suggest that the masking effects produced by the ipsilateral distracters were primarily energetic. Note that the number of errors contained in the distracter(s) never exceeds that expected by chance. Note also that at the lowest T/D ratios the number of errors contained in the ipsilateral distracter was lower than would be expected by chance. This suggests that at this T/D ratio, although the target was probably inaudible, the listener could hear and understand the ipsilateral distracter and thus eliminate it from the set of possible responses. The same strategy was apparently not used for the contralateral distracter, since the number of responses contained in this message was never different from chance.

Figure 10 shows the error analysis of the data from the oldest group of children (13.6–16). There are only slight differences between the error patterns shown here and those from the adults (Fig. 9). The same can be said of the error patterns for the children whose errors are not shown (ages 6.6–13.5). However, one minor difference emerges in the error data from the 6.6 to 8.5 year group. This group does not show a lower than chance frequency of errors contained in the ipsilateral distracter at low T/D ratios. In the adults and other groups of children this is interpreted as evidence of the ability to eliminate the ipsilateral distracter message from the set of response alternatives.

Figure 11 shows an error analysis for the 4.6–5.7 year old age group in the female distracter conditions. Note that for the monaural and contralateral noise distracter conditions (upper panels), the overwhelming majority of errors came from the distracter phrases, suggestive of informational masking. Since it is probably safe to assume that at the positive T/D ratios represented here both target and distracter were audible and understandable, one must conclude that the children simply could not determine or remember which was the target phrase and which was the distracter or could not disentangle the two. In the error analysis from the contralateral female speech distracter condition (lower panels), an especially intriguing result can be seen. As in the monaural condition, most of the errors in this condition were contained in the distracter(s). However, in contrast with the data from adults and older children, a large number of errors were contained in the contralateral distracter phrase. This could mean that young children forgot or could not determine which was the target voice, and, in addition, could not segregate the two messages in the ipsilateral ear. Thus, their only remaining choice was the contralateral distracter. This is qualitatively consistent with previous results from our laboratory (Wightman *et al.*, 2003) that indicated an inability of preschoolers to focus their attention on a single ear. The Wightman *et al.* (2003) experiments involved detection of a pure-tone signal which was presented with a random multicomponent tonal distracter. Children in the preschool age group demonstrated large amounts of informational masking when the distracter was in the target ear and only a modest release from masking when the distracter was presented to the nontarget ear.

The results from the female distracter conditions agree with those from other studies in the literature on speech-speech masking in which different sex target and distracter talkers were used. For example, in the experiment described by Doyle (1973), children (8, 11, and 14 years old) shadowed a male target talker in the presence of a single female distracter. Intrusions (errors) from the distracter talker were much more frequent in the younger children than in the older children, a result that appears consistent with ours. However, our results do not appear consistent with those from the recent study reported by Litovsky (2005). In that experiment children and adults identified target spondees produced by a male talker with either sentences (female talker) or modulated speech-spectrum noise used as distracters. Recognition thresholds were higher with the noise distracter for both children and adults, and there was no difference in amount of masking between children and adults. Thus, for reasons that are not obvious, it seems likely that informational masking was not a factor in the Litovsky (2005) study. The T/D ratios reported as thresholds in the Litovsky (2005) study are very low (−16 dB or poorer) compared to most of those tested here. Brungart (2001b) has shown that at such low T/D ratios, energetic masking exceeds informational masking in the case of different sex target and distracter talkers.

IV. CONCLUSIONS

Auditory selective attention was measured in 38 children (ages 4–16 years) and 8 adults using a closed-set speech-recognition task. The results suggested that listener performance was dominated by informational masking when the distracter was speech from a talker of the same sex as the target talker, and they confirmed earlier findings that children are much more influenced by informational masking than are adults. For the youngest children informational masking with the single ipsilateral distracter was more than 15 dB greater than in adults. A clear and monotonic age effect was also shown, with some children as old as 16 years still not performing at adult levels. For older children and adults, the added informational masking produced by a contralateral distracter amounted to about a 5 dB shift in the psychometric function toward poorer performance. For the youngest children the effect was somewhat smaller. Using a male target talker and female distracter(s) produced a considerable release from informational masking for all age groups tested.

Acknowledgments

The authors would like to thank Michael Callahan, Jen Junion Dienger, and the teachers at the University of Wisconsin Waisman Early Childhood Program for their contributions to the research. The research was supported financially by a grant to F.L.W. from the National Institutes of Health (Grant No. R01-HD023333-14).

References

Allen, P. Doctral dissertation. University of Wisconsin-Madison; 1991. Children's detection of auditory signals.

Allen P, Nelles J. Development of auditory information integration abilities. J Acoust Soc Am 1996;100:1043–1051. [PubMed: 8759957]

Allen P, Wightman F. Spectral pattern discrimination by children. J Speech Hear Res 1992;35:222–233. [PubMed: 1735972]

Allen P, Wightman F. Effects of signal and masker uncertainty on children's detection. J Speech Hear Res 1995;38:503–511. [PubMed: 7596115]

Allen P, Wightman F, Kistler D, Dolan T. Frequency resolution in children. J Speech Hear Res 1989;32:317–322. [PubMed: 2739383]

Bargones JY, Werner LA. Adults listen selectively; Infants do not. Psychol Sci 1994;5:170–174.

Bargones JY, Werner LA, Marean GC. Infant psychometric functions for detection: Mechanisms of immature sensitivity. J Acoust Soc Am 1995;98:99–111. [PubMed: 7608411]

Berman S, Friedman D. The development of selective attention as reflected by event-related brain potentials. J Exp Child Psychol 1995;59:1–31. [PubMed: 7876768]

Bolia RS, Nelson WT, Ericson MA, Simpson BD. A speech corpus for multitalker communications research. J Acoust Soc Am 2000;107:1065–1066. [PubMed: 10687719]

Bregman, AS. Auditory Scene Analysis. MIT; Cambridge, MA: 1990.

Brungart DS. Evaluation of speech intelligibility with the coordinate response measure. J Acoust Soc Am 2001a;109:2276–2279. [PubMed: 11386582]

Brungart DS. Informational and energetic masking effects in the perception of two simultaneous talkers. J Acoust Soc Am 2001b;109:1101–1109. [PubMed: 11303924]

Brungart DS, Simpson BD. Within-ear and across-ear interference in a cocktail-party listening task. J Acoust Soc Am 2002;112:2985–2995. [PubMed: 12509020]

Brungart DS, Simpson BD. Within-ear and across-ear interference in a dichotic cocktail party listening task: Effects of masker uncertainty. J Acoust Soc Am 2004;115:301–310. [PubMed: 14759023]

Brungart DS, Simpson BD, Ericson MA, Scott KR. Informational and energetic masking effects in the perception of multiple simultaneous talkers. J Acoust Soc Am 2001;110:2527–2538. [PubMed: 11757942]

Bryden MP, Allard FA. Do auditory perceptual asymmetries develop? Cortex 1981;17:313–318. [PubMed: 7285602]

Coch D, Sanders LD, Neville HJ. An event-related potential study of selective auditory attention in children and adults. J Cogn Neurosci 2005;17:605–622. [PubMed: 15829081]

Dirks DD, Bower DR. Masking effects of speech competing messages. J Speech Hear Res 1969;12:229–245. [PubMed: 5808851]

Doyle AB. Listening to distraction: A developmental study of selective attention. J Exp Child Psychol 1973;15:100–115. [PubMed: 4706960]

Egan JP, Carterette EC, Thwing EJ. Some factors affecting multi-channel listening. J Acoust Soc Am 1954;26:774–782.

Fallon M, Trehub SE, Schneider BA. Children's perception of speech in multitalker babble. J Acoust Soc Am 2000;108:3023–3029. [PubMed: 11144594]

Geffen G. Development of hemispheric specialization for speech perception. Cortex 1976;12:337–346. [PubMed: 1009770]

Geffen G. The development of the right ear advantage in dichotic listening with focused attention. Cortex 1978;14:169–177. [PubMed: 679698]

Geffen G, Sexton MA. The development of auditory strategies of attention. Dev Psychol 1978;14:11–17.

Gomes H, Molholm S, Christodoulou C, Ritter W, Cowan N. The development of auditory attention in children. Front Biosci 2000;5:D108–D120. [PubMed: 10702373]

Hall JW III, Grose JH, Buss E, Dev MB. Spondee recognition in a two-talker masker and a speech-shaped noise masker in adults and children. Ear Hear 2002;23:159–165. [PubMed: 11951851]

Hillyard SA, Hink RF, Schwent VL, Picton TW. Electrical signs of selective attention in the human brain. Science 1973;182:177–180. [PubMed: 4730062]

Hiscock M, Kinsbourne M. Selective listening asymmetry in preschool children. Dev Psychol 1977;13:217–224.

Hiscock M, Kinsbourne M. Asymmetries of selective listening and attention switching in children. Dev Psychol 1980;16:70–82.

Hugdahl K, Carlsson G, Eichele T. Age effects in dichotic listening to consonant-vowel syllables: Interactions with attention. Dev Neuropsychol 2001;20:445–457. [PubMed: 11827098]

Kidd G Jr, Mason CR, Arbogast TL, Brungart DS, Simpson BD. Informational masking caused by contralateral stimulation. J Acoust Soc Am 2003;113:1594–1603. [PubMed: 12656394]

Litovsky R. Speech intelligibility and spatial release from masking in young children. J Acoust Soc Am 2005;117:3091–3099. [PubMed: 15957777]

Lutfi RA. Informational processing of complex sound. I. Intensity discrimination. J Acoust Soc Am 1989;86:934–944. [PubMed: 2794246]

Lutfi RA. A model of auditory pattern analysis based on component-relative-entropy. J Acoust Soc Am 1993;94:748–758. [PubMed: 8370881]

Lutfi RA, Kistler DJ, Callahan MR, Wightman FL. Psychometric functions for informational masking. J Acoust Soc Am 2003a;114:3273–3282. [PubMed: 14714808]

Lutfi RA, Kistler DJ, Oh EL, Wightman FL, Callahan MR. One factor underlies individual differences in auditory informational masking within and across age groups. Percept Psychophys 2003b;65:396–406. [PubMed: 12785070]

Maccoby EE. The development of stimulus selection. Minnesota Symposia on Child Psychology 1969;3:68–96.

Maccoby EE, Konrad KW. Age trends in selective listening. J Exp Child Psychol 1966;3:113–122. [PubMed: 5939103]

Morris R, Bakker D, Satz P, Van der Vlugt H. Dichotic listening ear asymmetry: Patterns of longitudinal development. Brain Lang 1984;22:49–66. [PubMed: 6722528]

Neff DL, Callaghan BP. Effective properties of multicomponent simultaneous maskers under conditions of uncertainty. J Acoust Soc Am 1988;83:1833–1838. [PubMed: 3403798]

Oh EL, Lutfi RA. Nonmonotonicity of informational masking. J Acoust Soc Am 1998;104:3489–3499. [PubMed: 9857508]

Oh EL, Wightman F, Lutfi RA. Children's detection of pure-tone signals with random multitone maskers. J Acoust Soc Am 2001;109:2888–2895. [PubMed: 11425131]

Pohl P, Grubmuller HG, Grubmuller R. Developmental changes in dichotic right ear advantage (REA). Neuropediatrics 1984;15:139–144. [PubMed: 6483111]

Richards VM, Tang Z, Kidd GD Jr. Informational masking with small set sizes. J Acoust Soc Am 2002;111:1359–1366. [PubMed: 11931313]

Schneider BA, Trehub SE, Morrongiello BA, Thorpe LA. Developmental changes in masked thresholds. J Acoust Soc Am 1989;86:1733–1742. [PubMed: 2808922]

Sexton MA, Geffen G. Development of three strategies of attention in dichotic monitoring. Dev Psychol 1979;15:299–310.

Stellmack MA, Willihnganz MS, Wightman FL, Lutfi RA. Spectral weights in level discrimination by preschool children: Analytic listening conditions. J Acoust Soc Am 1997;101:2811–2821. [PubMed: 9165735]

Tang Z, Richards VM. Examination of a linear model in an informational masking study. J Acoust Soc Am 2003;114:361–367. [PubMed: 12880047]

Wichmann FA, Hill NJ. The psychometric function. I. Fitting, sampling, and goodness of fit. Percept Psychophys 2001a;63:1293–1313. [PubMed: 11800458]

Wichmann FA, Hill NJ. The psychometric function. II. Bootstrap-based confidence intervals and sampling. Percept Psychophys 2001b;63:1314–1329. [PubMed: 11800459]

Wightman, FL.; Allen, P. Individual differences in auditory capability among preschool children. In: Werner, LA.; Rubel, EW., editors. Developmental Psychoacoustics. American Psychological Association; Washington, DC: 1992. p. 113-133.

Wightman FL, Callahan MR, Lutfi RA, Kistler DJ, Oh E. Children's detection of pure-tone signals: Informational masking with contralateral maskers. J Acoust Soc Am 2003;113:3297–3305. [PubMed: 12822802]

Willihnganz MS, Stellmack MA, Lutfi RA, Wightman FL. Spectral weights in level discrimination by preschool children: Synthetic listening conditions. J Acoust Soc Am 1997;101:2803–2810. [PubMed: 9165734]

Wright BA, Saberi K. Strategies used to detect auditory signals in small sets of random maskers. J Acoust Soc Am 1999;105:1765–1775. [PubMed: 10089600]

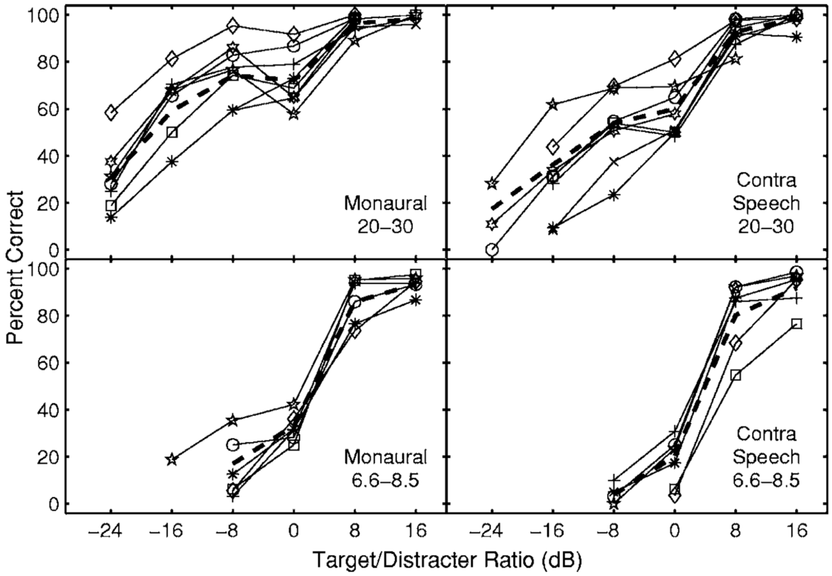

FIG. 1.

Psychometric functions from individual listeners in two age groups (adults—top panels; children 6.6–8.5 years—bottom panels) and two male distracter conditions (monaural—left panels; contralateral—right panels). Functions plotted with different symbols represent data from the different listeners and the dashed line represents the mean. Data are plotted as a function of the target/distracter ratio in the target ear.

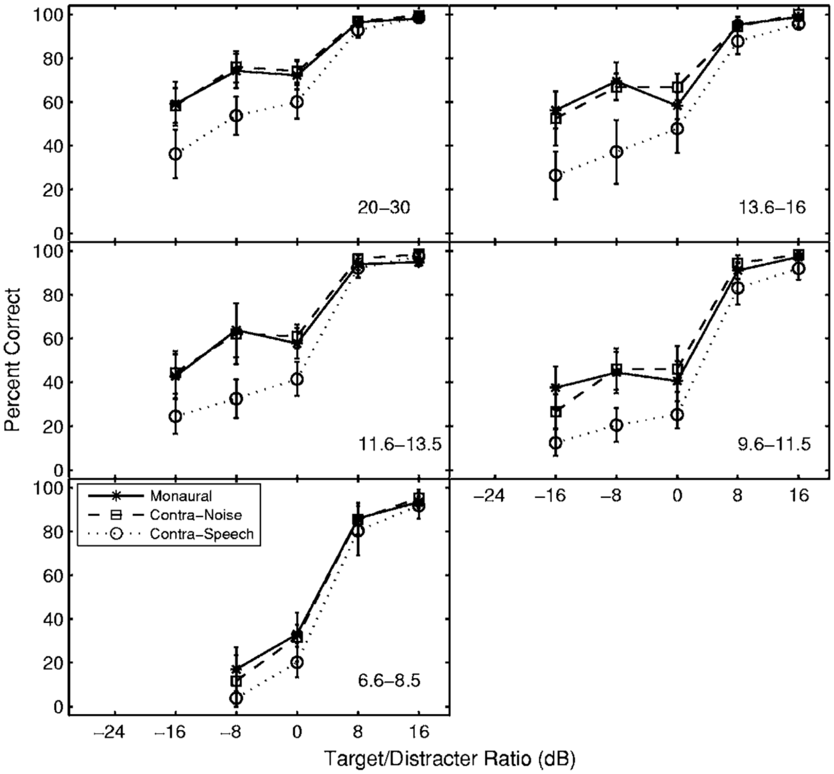

FIG. 2.

Mean psychometric functions for listeners in five age groups in the three conditions that used a male distracter talker. The age ranges are indicated in each panel and the conditions are indicated in the lower left panel. Error bars represent 95% confidence intervals of the mean.

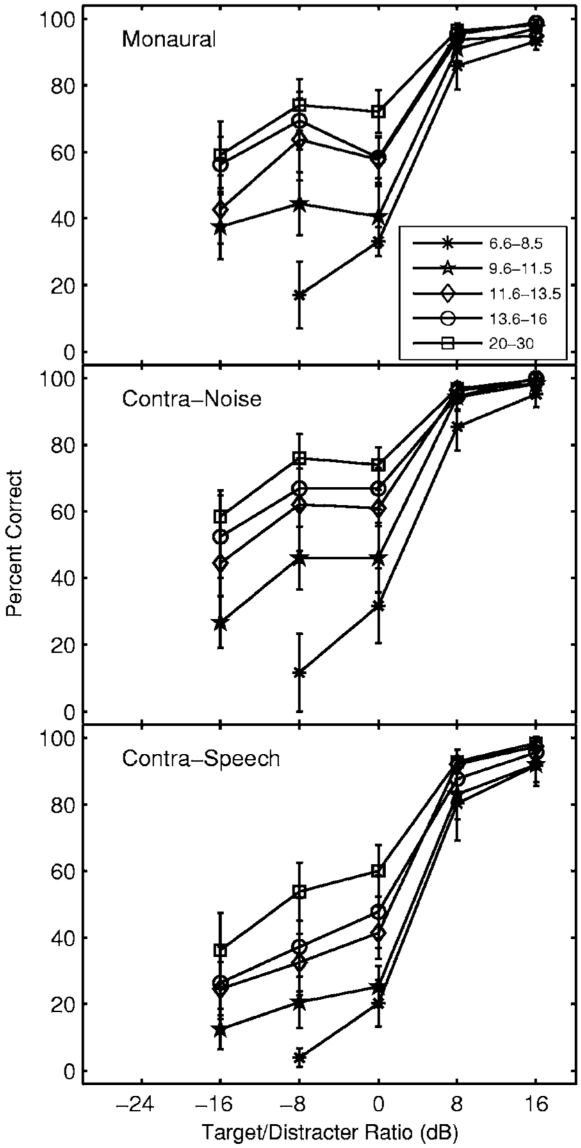

FIG. 3.

Mean psychometric functions from the three male distracter conditions plotted with age group as the parameter in each panel. Error bars represent 95% confidence intervals of the mean.

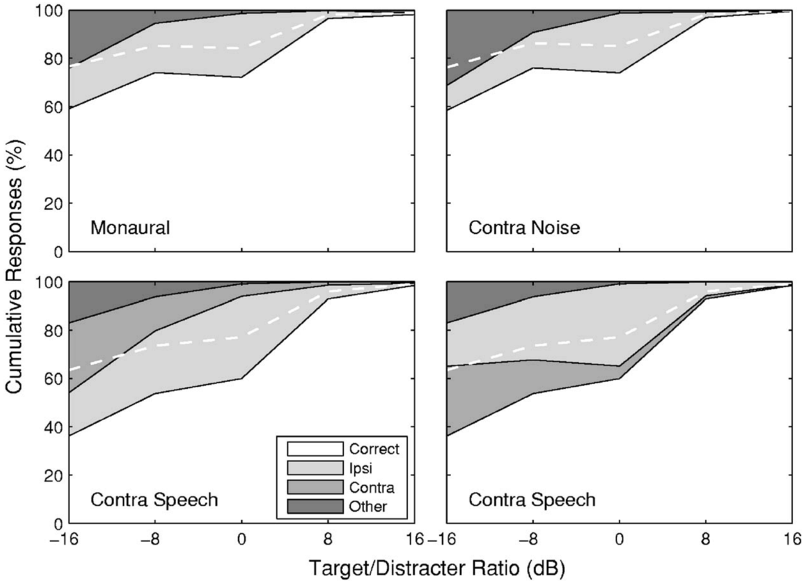

FIG. 4.

Response analysis of the data from the adult listeners in the three male masker conditions. The upper panels show responses from the monaural condition (left) and the contralateral noise condition (right). Lower panels show responses from the contralateral speech distracter condition. Correct responses are indicated by the lightest shading. Color and/or number errors contained in the ipsilateral distracter, contralateral distracter, or neither are indicated by darker shadings. The light dashed lines indicate the number of color and/or number errors expected if the listener were responding randomly. The same data are plotted in the two lower panels to facilitate comparison of actual responses to expected random responses from the ipsilateral (left) and contralateral (right) distracters.

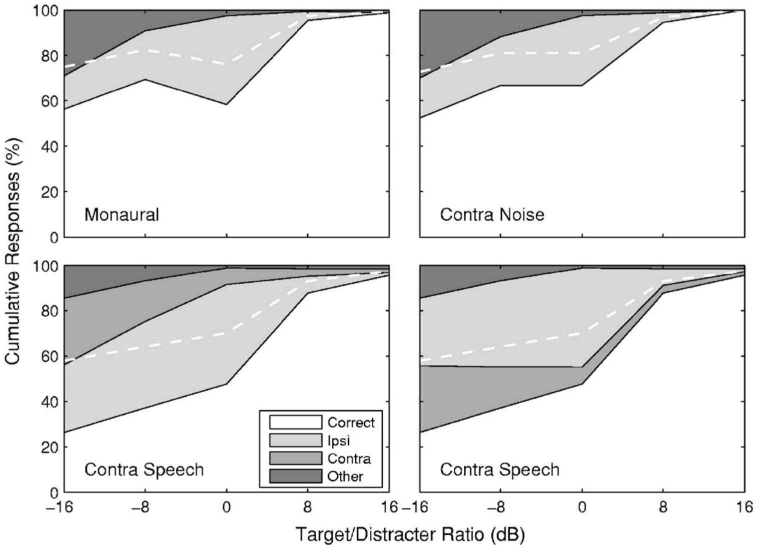

FIG. 5.

Same as Fig. 4, except that the response analysis of the data from the children in the 13.6–16.0 year age group is shown.

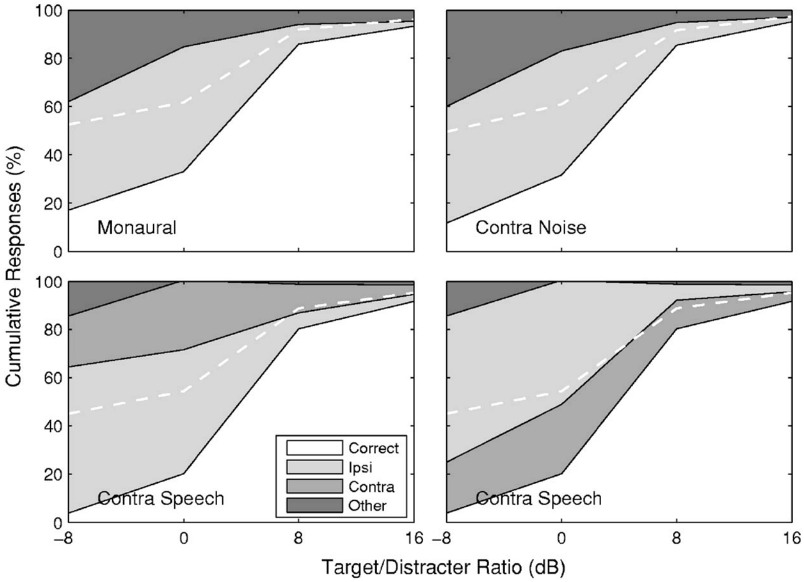

FIG. 6.
Same as Fig. 4, except that the response analysis of the data from the children in the 6.6–8.5 year age group is shown.

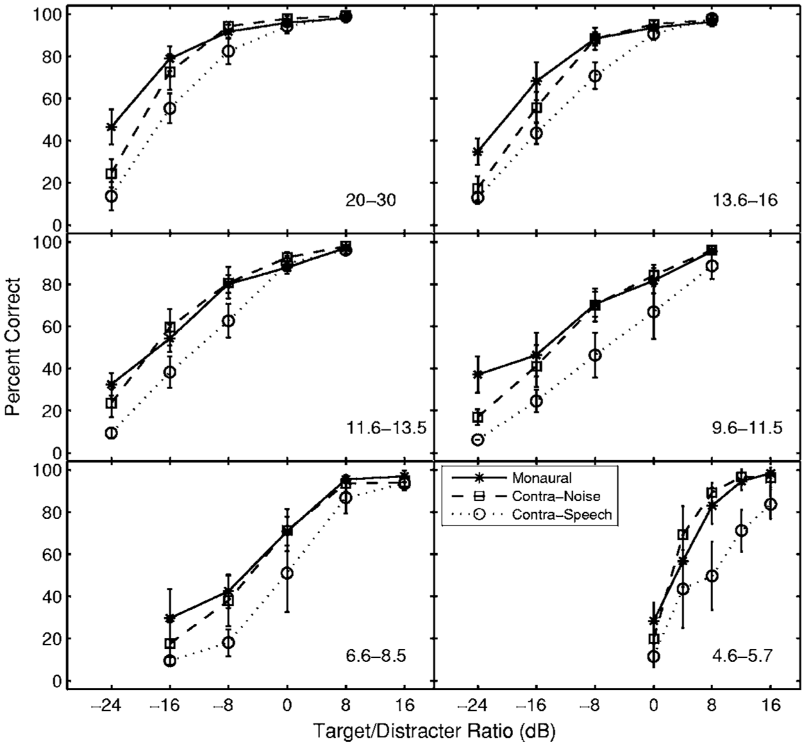

FIG. 7.

Mean psychometric functions from listeners in all age groups from the conditions that used the female distracter talker. Labeling is the same as in Fig. 2. Error bars represent 95% confidence intervals of the mean.

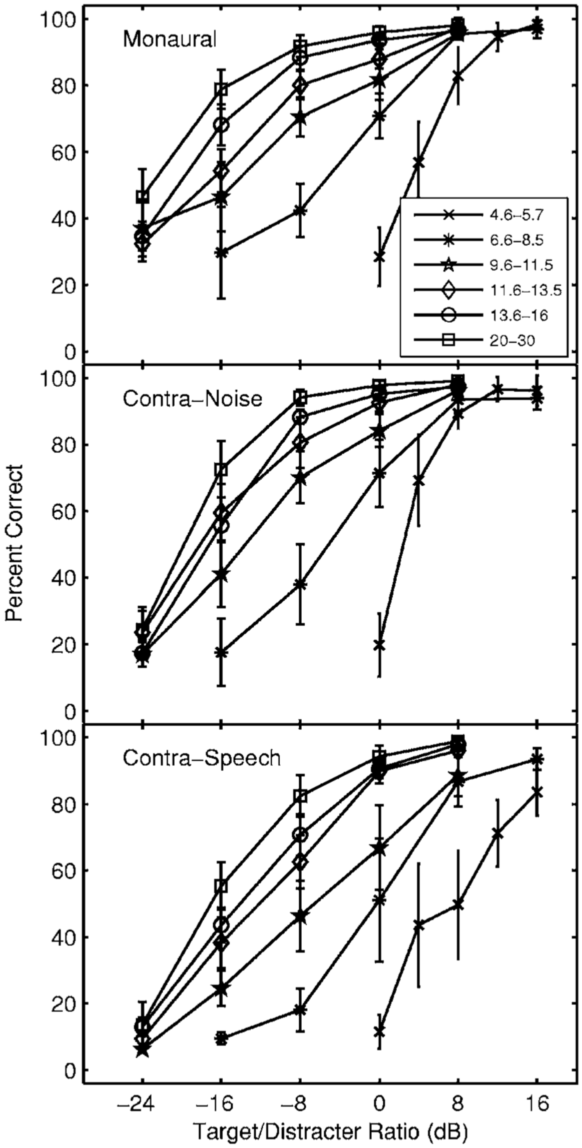

FIG. 8.

Same as Fig. 3 except for the female distracter conditions. Error bars represent 95% confidence intervals of the mean.

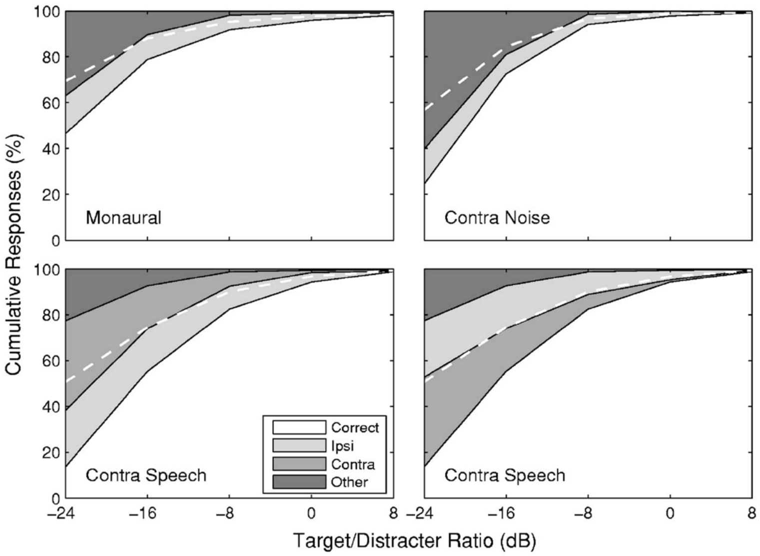

FIG. 9.

Same as Fig. 4 except that the response analysis of the data from the adults in the female distracter conditions is shown.

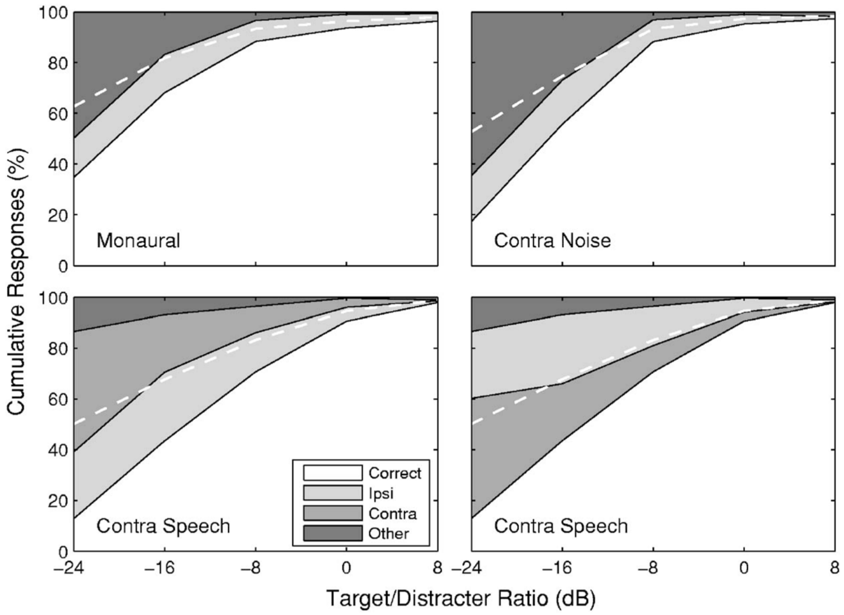

FIG. 10.

Same as Fig. 4 except that the response analysis of the data from the children in the 13.6–16.0 year age group in the female distracter conditions is shown.

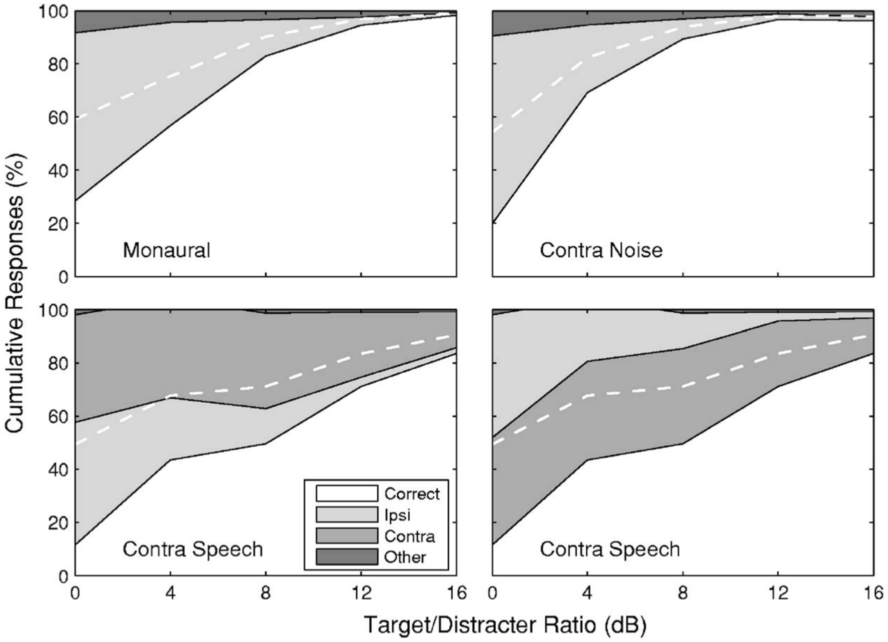

FIG. 11.

Same as Fig. 4 except that the response analysis of the data from the children in the 4.6–5.7 year age group in the female distracter conditions is shown. Note that the abscissa represents different T/D ratios than in the other figures.