Infants' detection and discrimination of sounds in modulated maskers

Lynne A. Werner

Department of Speech and Hearing Sciences, University of Washington, Seattle, Washington 98105-6246

Adults and 7-month-old infants were compared in detection and discrimination of sounds in modulated maskers. In two experiments, the level of a target sound was varied to equate listeners' performance in unmodulated noise, and performance was assessed at that level in a noise modulated with the envelope of single-talker speech. While adults' vowel discrimination and tone detection were better in the modulated than in the unmodulated masker, infants' vowel discrimination was poorer in the modulated than in the unmodulated masker. Infants' tone detection was the same in the two maskers. In two additional experiments, each age group was tested at one level with order of testing in modulated and unmodulated maskers counterbalanced across subjects. Both infants and adults discriminated between vowels better in single-talker modulated and sinusoidally amplitude modulated (SAM) maskers than in an unmodulated masker, but infants' modulated-unmodulated difference was smaller than than that of adults. Increasing the modulation depth of the SAM masker did not affect the size of infants' modulated-unmodulated difference. However, infants' asymptotic performance in a modulated masker limits the extent to which their performance could improve. Infants can make use of information in masker dips, but masker modulation may also interfere with their ability to process the target.
© 2013 Acoustical Society of America. [http://dx.doi.org/10.1121/1.4803903]

I. INTRODUCTION

It is well established that adults detect sounds and perceive speech better in a modulated masker than in an unmodulated masker (e.g., Bacon and Lee, 1997; Buss *et al.*, 2003; Howard-Jones and Rosen, 1993). The common explanation for this effect is that listeners give greater weight to inputs at minima in the masker waveform, when the target-to-masker ratio (TMR) is better, than at other times, allowing the target sound to be detected or identified at lower intensities. This listening strategy is referred to as "dip listening."

Many factors influence the modulated-unmodulated difference (MUD), including the frequency and modulation spectra of the signal and masker (e.g., Bacon and Lee, 1997; Bacon *et al.*, 1997; Kwon and Turner, 2001; Oxenham and Simonson, 2009). Listener characteristics are also important factors: Hearing-impaired and older listeners may not demonstrate a MUD (e.g., Bernstein and Grant, 2009; Dubno *et al.*, 2002; Eisenberg *et al.*, 1995; Festen and Plomp, 1990; Nelson *et al.*, 2003; Peters *et al.*, 1998; Summers and Molis, 2004; Takahashi and Bacon, 1992). A listener's ability to follow modulation—temporal resolution—will clearly influence the degree to which masker modulation improves target processing. In addition, listening in the dips implies the ability to differentially weight inputs across time: If a listener bases decisions about the target on input averaged over periods including both peaks and valleys in the modulated masker, little or no MUD would result. When the target is a word or a sentence, it is also important that the listener be able to reconstruct the target from incomplete information, known as "glimpsing."

The development of the MUD has not been extensively studied. Grose *et al.* (1993) measured children's detection thresholds for 500 and 2000 Hz tones in modulated and unmodulated noise. The MUD was greater for adults than for 4- to 10-yr-old children. Although the MUD appeared to increase with age between 4 and 10 yr, the age effect was not significant across that age range. Recently Hall *et al.* (2012) compared speech reception thresholds of children and adults in steady speech spectrum noise to speech reception thresholds in noise that was modulated temporally, spectrally or both spectrally and temporally. They found that both younger (∼6 yr old) and older (∼9 yr old) children and adults had equivalent MUD for spectrally modulated maskers. The younger children, however, had smaller MUD than adults for temporally and spectrally temporally modulated maskers; the older children had smaller MUD than adults only in spectrally temporally modulated maskers. Stuart (2005, 2008) examined school-aged children's speech recognition in interrupted noise but with less conclusive results. Because the temporal modulation transfer function (TMTF) of children as young as 4 yr of age is adult like in shape (Hall and Grose, 1994), it appears that the reduced MUD among child listeners does not result from immature temporal resolution.

On the basis of these observations of children, one might predict that infants would also have a smaller MUD than adults. A recent study by Newman (2009) investigated this issue by determining how the number of competing talkers influenced infants' ability to recognize their own name. When there is only one competing talker, the fluctuations in competing speech will be relatively large; when there are many competing talkers, the fluctuations will be small. Normal-hearing adults' speech recognition is better with one or two competing talkers than with a greater number of competing talkers (e.g., Drullman and Bronkhorst, 2004; Freyman *et al.* 2004). Newman compared 5- and 8.5-month-old infants' recognition

of their own name when the competing sound was single-talker speech, nine-talker speech, and time-reversed single-talker speech. Both groups of infants recognized their own names at 10-dB target-to-masker ratio when the competing sound was nine-talker speech but not when it was normal or time-reversed single-talker speech.

As Newman (2009) pointed out, there are several explanations for infants' poor performance with a single-talker masker. The first is that immature temporal resolution prevented infants from taking advantage of the fluctuations in single-talker speech. Infants have poor gap detection thresholds (Trehub *et al.*, 1995; Werner *et al.*, 1992), but by 6 months, they are no more susceptible to forward masking than adults are (Werner, 1999). Although infants require greater modulation depth than adults do to detect amplitude modulation, the similarity in the effect of modulation frequency on modulation detection by infants and adults suggests that temporal resolution is close to mature as early as 3 months of age (Werner, 2006). A second explanation is related to the fact that infants detect, discriminate, or recognize a target at higher TMR compared to adults. At higher TMR adults' MUD is also reduced (e.g., Bernstein and Grant, 2009; Oxenham and Simonson, 2009), and there is some evidence that this effect can account for at least part of the difference between adults' and young children's MUD (Hall *et al.*, 2012). Another explanation is that infants are inefficient at dip listening. This hypothesis is consistent with the previously cited results for older children (Grose *et al.*, 1993; Hall *et al.*, 2012). A fourth explanation is based on the fact that when the competing sound is speech, the spectral-temporal variation in the masker is likely to cause informational masking (Brungart and Simpson, 2007). In adults, informational masking may offset the advantage of masker fluctuations in speech recognition for two- and three-talker maskers (e.g., Freyman *et al.*, 2004). Although single-talker speech was the masker in Newman's study, infants are known to be more susceptible than adults to informational masking, or at least to masking by sounds that are similar to the target (Leibold and Werner, 2006). Finally, infants may have difficulty recognizing the target on the basis of partial information (e.g., Fernald *et al.*, 2001; but see Johnson, 2004).

The experiments reported here further investigated infants' ability to take advantage of masker fluctuations in discrimination and detection, using noise maskers, which lack spectral variation and are less similar to the target than a speech masker. Two modulated maskers were examined, noise modulated with the envelope of single-talker speech and sinusoidally amplitude modulated noise. While the single-talker modulated noise allows comparison to Newman's (2009) previous results, the modulation spectrum of sinusoidally amplitude modulated (SAM) noise could be more easily manipulated to examine stimulus factors contributing to the MUD, such as modulation rate and depth. The fluctuations in SAM masker are also more regular and predictable. In all but one experiment, the listeners' task was to detect a change in a repeating vowel, from /a/ to /i/ or from /i/ to /a/. This task is somewhat simpler than that employed by Newman; it might be expected to be less influenced by the listener's ability to identify the target on the basis of partial information. Finally, the test method used provided an estimate of the MUD for individual subjects and allowed direct comparison of infants' and adults' MUD.

II. EXPERIMENT 1: VOWEL DISCRIMINATION IN SINGLE-TALKER MODULATED NOISE

A. Methods

1. Stimuli

The target sounds were single exemplars of the vowels /a/ and /i/. The vowels were synthesized using SYNTHWORKS, a Klatt-based speech synthesizer program. The formant frequencies of each vowel were set to the average formant frequencies of six male native speakers of English from the Pacific Northwest (Bor *et al.*, 2008). Each vowel was 200 ms in duration, presented at 800 ms intervals. The maskers came from the Connected Speech Test (CST, Cox *et al.*, 1987). The noise had the long-term spectrum of the CST sentences. The single-talker envelope was constructed by digitally concatenating the 12 CST sentences, then rectifying and low-pass filtering the waveform at 30 Hz. A 30-s segment of the envelope was multiplied by the CST-spectrum noise to create the single-talker modulated (ST) masker. The unmodulated noise (UN) masker was the same 30-s segment of CST-spectrum noise with no additional modulation. The 30-s noise segment was repeated throughout the test session without interruption. The RMS amplitudes of the masker waveforms were equated. The level of the maskers was 60 dB sound pressure level (SPL). The stimuli were presented via ER-1 insert phones. Child-size or trimmed adult-size foam ear tips were used for infant listeners. A Zwislocki coupler was used for calibration. An in-the-ear calibration check was made at the beginning of each test session and at any time that the ear tip was replaced during testing using an ER-7 C microphone. Participants were tested in a sound booth.

2. Subjects

Normal hearing infants and young adults participated. The 12 infants who provided a complete set of data had an average age of 30.6 weeks (SD = 3.5). By parental report, all infant participants were developing normally, were healthy, passed newborn hearing screening, had no identified hearing loss, and had no risk factors for hearing loss. Infants had two or fewer episodes of otitis media and treatment for any episode of otitis media was completed at least 1 wk prior to testing. Ten adults provided data; their average age was 22.1 yr (SD = 2.5). By self-report, no adult participants had identified hearing loss, risk factors for hearing loss, or history of chronic otitis media. Adult participants had no more than 2 yr of musical training. All participants passed tympanometry with 220 and 1000 Hz probes on each test day. Testing was attempted with 13 infants who never reached training criteria, 14 infants who met training criteria but never completed a block of test trials, and 8 infants who completed only one condition after three to four test

sessions. Testing was attempted with three additional adults who completed only one condition in their test session.

3. Procedure

a. Overview. In each test session, the participant listened to a vowel that repeated at 800 ms intervals, presented in a noise background. The listener's task was to respond when the vowel changed from /a/ to /i/ or from /i/ to /a/ on one presentation. Both "change" and "no-change" trials were included.

The listener was tested first with the UN masker to identify the vowel level that produced a $d' \sim 1$, so that the effect of modulation could be assessed from the same baseline sensitivity for all listeners. The initial level of the vowels was set to a value at which the listener was expected to achieve a $d' \sim 1$ in detecting the vowel change in UN noise based on pilot testing. The starting level was 56 dB SPL for infants and 46 dB SPL for adults. Using a binomial model, it was determined that if the infant/observer team is guessing on every trial, with 15 change and 15 no-change trials, a d' greater than 0.8 will be achieved less than 5% of the time. If d' was higher than 1.4 or lower than 0.8, the level of the vowels was adjusted and testing was repeated. Once a level producing a $d' \sim 1$ in UN was identified, the listener was tested again with the vowels set at that level, but in the ST masker. The average final test level for the infants was 57.0 (SD = 1.0) and for the adults was 45.6 (SD = 1.0) dB SPL.

b. Observer-based procedure. An observer-based procedure (Werner, 1995) was used to assess infants' sensitivity to the vowel change. Infants' were seated on a caregiver's lap, facing a window and a video camera. An assistant seated to the infants' left manipulated quiet toys to keep the infant facing forward, calm, and vaguely entertained. Both the assistant and caregiver wore headphones to prevent them from hearing any of the sounds presented to the infant. The caregiver listened to music; the assistant listened to an observer outside the booth. As soon as the infant's insert ear tip had been positioned and the occupants of the booth were settled in, the observer turned on the noise and the repeating background vowel. The sounds were presented throughout the session. The background vowel was /a/ for half of the subjects and /i/ for half of the subjects.

The observer watched the infant and the adults in the booth. When the infant was quiet and attentive, the observer signaled the computer to begin a trial. One of two types of trials then occurred, chosen randomly. On a change trial, the target vowel was presented once instead of the background vowel. On a no-change trial, the background vowel was presented again. The observer did not know which type of trial would occur. On each trial, the observer judged whether or not a change trial had occurred within 4 s of trial onset. The observer received feedback after each trial. Because the infant's behavior provided the only basis for this judgment, if the observer could reliably distinguish change from no-change trials, then the infant must have been able to discriminate the vowels. To ensure that the infant would respond to a vowel change, an interesting visual display—a mechanical toy or a video—was turned on as soon as the observer correctly identified a change trial. The interesting visual display is referred to as a "visual reinforcer."

To teach the infant to respond to vowel changes and to allow the observer to determine how the infant responded to vowel changes, each infant completed a training procedure. Initially, the level of the vowels was set at 66 dB SPL, expected to be clearly audible, and 80% of the trials were change trials. The visual reinforcer was activated as soon as the observer correctly identified a change trial or after 4 s if the observer missed a change trial. This training phase demonstrated the association between the vowel change and the visual reinforcer to the infant. Once the observer reached a criterion of four of five correct change trials with at least one correct no-change trial, the contingencies changed. The probability of a change trial was 0.5, and the visual reinforcer was only activated if the observer correctly identified a change trial within the 4-s window. This phase continued until the observer reached a criterion of at least four of the last five change trials and at least four of the last five no-change trials correct. For infants, the average number of trials required to reach criterion was 20.4 (SD = 18.9).

Once training criteria had been met, the level of the vowels was reduced to the test level, as described in the preceding text. A block of 15 change and 15 no-change trials, randomly ordered, was completed. Five "probe" trials were randomly interspersed in the test block; on probe trials the level of the changed vowel was presented at the training level. The probe trials were used to determine whether or not the infant was still on task. A test block was only included in the final analyses if the observer correctly identified at least three of the five probe trials.

The session was terminated if the infant became fussy or sleepy or if all trials in a condition were completed. If the session was terminated before all the trials in a test block were completed, a few reminder training trials were completed at the beginning of the next session, and the interrupted test block was replaced. Infants were scheduled for two to four test sessions. All infants required at least two sessions to complete both masker conditions; most infants required three sessions to do so.

Adults were tested using the same basic procedure. Adults listened alone in the sound booth. They were instructed to raise a hand when they "heard the sound that makes the mechanical toy come on." An observer recorded the adult's response. The level of the vowels in training was 56 dB SPL; otherwise the same training procedures and criteria were applied. The average number of training trials required to meet criterion was 10.3 (SD = 5.5) for adults. All adults completed both the UN and ST masker conditions in one session.

Performance was described in terms of d' where $d' = Z[p(\text{hit})] - Z[p(\text{false alarm})]$ for each individual subject. If the hit rate was 1.0 or the false alarm rate was 0, d' was calculated after adjusting the hit rate or false alarm rate by the reciprocal of the number of trials.

B. Results and discussion

Individual d' for infants and adults in the UN and ST masker conditions is plotted in Fig. 1; mean d' is indicated by the horizontal lines. As expected, mean d' was higher for the ST masker than for the UN masker for adults. Of the
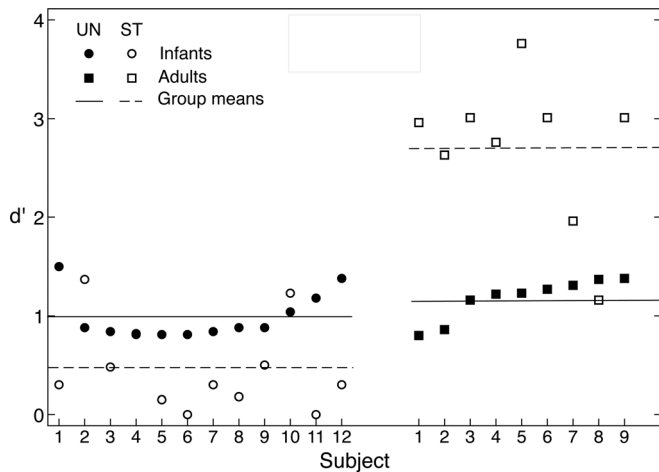
FIG. 1. d' in vowel discrimination for individual subjects for single-talker (ST) modulated and unmodulated (UN) noise maskers from experiment 1. Level of the vowels was adjusted for each subject individually to equate performance in UN noise masker condition. Solid line represents mean d' in UN noise masker and dashed line represents mean d' in the ST modulated noise masker by age group.

nine adults tested, eight had higher d' for the ST masker than for the UN masker. Because one adult subject performed perfectly in the ST noise condition, the mean benefit of modulation for the adults is likely underestimated at least slightly. The results for infants were different: d' for the ST masker was actually lower than for the UN masker. Only 2 of the 12 infants had higher d' for the ST masker than for the UN masker; 9 of the 12 infants had higher d' for the UN masker than for the ST masker. An age group × masker analysis of variance (ANOVA) with repeated measures on Masker confirmed a significant Age group × masker interaction [$F(1,19) = 47.24$, $p < 0.0001$]. *Post hoc* paired t-tests, with Bonferroni correction, were used to test the effect of masker in the two age groups. For adults, d' was significantly higher for the ST than the UN masker [$t(8) = -5.54$, $p = 0.0003$]. For infants, d' was significantly lower for the ST than for the UN masker [$t(11) = 3.35$, $p = 0.0032$].

The results, then, showed the expected positive MUD for adults, but a negative MUD for the infants. There are several explanations for the infants' result. Experiment 2 tested the possibility that the deficit in vowel discrimination associated with the ST masker was peculiar to speech or to the necessity to monitor the ongoing stream of vowels throughout the session to identify a change.

III. EXPERIMENT 2: TONE DETECTION IN ST NOISE

The purpose of this experiment was to determine whether infants were able to take advantage of single-talker modulation in a masker to improve detection of a tone, a nonspeech target sound. In this task, infants would not be required to monitor an ongoing target sound stream.

A. Methods

1. Stimuli

The maskers were identical to those used in experiment 1. The target sound was a 1-kHz tone. The tone was 50 ms in

duration with 10-ms $\cos^2$ ramps. A relatively short duration tone was used because adults' detection of longer duration tones in the ST masker was perfect at levels that produced a d' of 1 in the UN masker. The starting level of the tone was chosen based on pilot testing, 58 dB SPL for infants and 44 dB SPL for adults. The average level of the tone producing a $d' \sim 1$ was 58.1 (SD = 0.8) dB SPL for the infants and 44.5 (SD = 1) dB SPL for the adults.

2. Subjects

All subjects met the same criteria for inclusion as in experiment 1. Data were obtained from 26 infants and 12 adults. The infants' age averaged 31.7 wk (SD = 3.9 wk). An additional four infants never met training criteria, and 16 infants completed only the UN masker condition. The adults' age averaged 23.4 yr (SD = 2.6). No adult data were excluded from the analysis.

3. Procedure

Essentially the same procedures were followed as in experiment 1. The masker was turned on at the beginning of the session. When the observer started a trial, either a tone trial or a no-tone trial occurred. The observer's task was to determine whether a tone trial had occurred based solely on the infant's behavior. The level of the tone in training and on probe trials was 65 dB SPL for infants and 55 dB SPL for adults. Infants required, on average, 29.1 trials (SD = 10.4) to reach training criteria, while adults required an average of 15.0 trials (SD = 1.7).

B. Results and discussion

Individual d' for the detection of the 1-kHz tone in the UN and ST maskers is plotted for infants and adults, with means indicated by the horizontal lines in Fig. 2. As in experiment 1, adults' mean d' was higher in the ST masker condition than in the UN masker condition. All 12 adults had a higher d' in the ST than in the UN masker condition. Again, the infants' results differed from those of the adults; however, in this case, infants achieved about the same d', on average, in
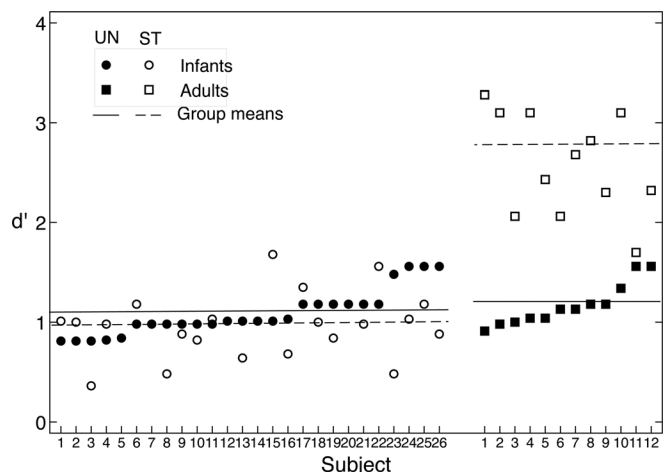

FIG. 2. d' in tone detection for individual subjects for ST modulated and UN noise maskers from experiment 2. Level of the vowels was adjusted for each subject individually to equate performance in UN noise masker condition. Solid line represents mean d' in UN noise masker and dashed line represents mean d' in the ST modulated noise masker by age group.

the two masker conditions. Of the 26 infants, 8 had a higher d' in the ST than in the UN masker condition and 13 had higher d' in the UN than in the ST masker condition.

An age group $\times$ masker ANOVA of d' with repeated measures on masker indicated significant effects of age group $[F(1,36) = 137.64, \quad p < 0.0001]$, masker $[F(1,36) = 63.28, p < 0.0001]$, and the age group $\times$ masker interaction $[F(1,36) = 90.04, \quad p < 0.0001]$. *Post hoc* paired *t*-tests showed that mean d' was significantly higher in the ST than in the UN masker condition for adults $[t(11) = -7.61, p < 0.0001]$ but that the difference was not significant for the infants $[t(25) = 1.76, p = 0.09]$.

The results of experiment 2, then, showed that even when the task was to detect a tone rather than to detect a change in an ongoing stream of vowels adults, but not infants, had a positive MUD. In the case of tone detection, however, infants' sensitivity was no poorer in the ST masker than in the UN masker.

A possible explanation for the results of both experiments 1 and 2, however, is that while the infants were capable of dip listening, the advantage of doing so was offset, or even more than offset, by negative effects of testing the modulated masker after the unmodulated masker. Testing the unmodulated masker first allowed the target level to be set appropriately for each individual. However, because it often required several test blocks to identify the target level that would produce a d' of 1, the ST masker condition was often tested in a third or even fourth visit to the lab. It is possible that the initial exposure to the unmodulated masker biased infants to listen in a particular way or that infants simply lost interest in the task. Experiment 3 was conducted to address this issue.

IV. EXPERIMENT 3: ORDER EFFECTS IN VOWEL DISCRIMINATION IN ST NOISE

The purpose of experiment 3 was to determine whether infants demonstrated a positive MUD in vowel discrimination when the order of testing ST and UN masker conditions was counterbalanced across subjects. In experiments 1 and 2, target levels were manipulated to equate performance across subjects in the UN noise masker condition prior to testing the ST noise masker condition. In experiment 3, all infants and all adults were tested in vowel discrimination at a level expected to produce a $d' \sim 1$ in the unmodulated masker condition, based on results of experiment 1.

A. Methods

1. Stimuli

The stimuli were identical to those used in experiment 1. For infants, the level of the vowels was set at 65 dB SPL in the training phase and on probe trials. In the test phase, one group of infants was tested at 56 dB SPL (1 dB lower than the average level in experiment 1) and at a second group of infants was tested at 58 dB SPL (1 dB higher than the average level in experiment 1). The 58 dB SPL group was added because many infants did not provide data at 56 dB SPL. For adults, the corresponding levels were 55 dB SPL in the training phase and 45 dB SPL in the test phase.

2. Subjects

Sixteen infants provided data for both the ST and UN noise masker conditions, 8 at each target level. Their average age was 29.7 wk (SD = 2.7). The inclusion criteria were the same as in experiments 1 and 2. Eight additional infants never reached training criteria; six of the eight infants in this category were tested first in the UN masker condition. Six infants reached training criterion but did not complete a test block; two of those six infants were tested first in the UN masker condition. Seven infants completed testing only in one masker condition; four of those infants completed testing only in the UN masker condition. Of the total 21 infants who did not provide data, 14 were tested with the 56 dB SPL vowels. That the success rate in this study was lower than that in experiments 1 and 2 is not unexpected because with a fixed target level in the test phase, the target could be below threshold for some infants. The average number of trials to criterion for infants was 28.3 (SD = 12.3) in the ST masker condition and 26.6 (SD = 10.2) in the UN masker condition. The average number of trials to criterion for adults was 9.7 (SD = 1.4) in the ST masker condition and 10.0 (SD = 2.0) in the UN masker condition.

3. Procedure

Each subject was randomly assigned to be tested first in the UN or ST masker condition. The subjects tested at 58 dB were tested after those at 56 dB SPL. If the subject completed all test trials, the test block was included in the final dataset. All other procedures were identical to those used in experiment 1.

B. Results and discussion

On the basis of preliminary analysis of the data of infants who completed both conditions, test order was not considered in subsequent analyses. Similarly, initial analyses of the infants' data indicated that the effect of vowel level was not statistically significant and that vowel level did not interact with masker condition. For that reason, the data for the infants tested at the two vowel levels were combined. In the ST masker condition, all 16 infants achieved a d' that was significantly greater than 0.8, which would be expected to be achieved by chance less than 5% of the time.

Individual d' is plotted for infants and adults for the two masker conditions in Fig. 3, with mean d' indicated by the horizontal lines. The infants had an average MUD ~ 0.2 d'. Individually, 12 of 16 infants had higher d' in the ST masker condition than in the UN masker condition. The adults had an average MUD of ~ 1.3 d', and seven of eight adults had higher d' in the ST masker condition than in the UN masker condition.

An age group $\times$ masker ANOVA of d' was conducted. The age group $\times$ masker interaction was significant $[F(1,24) = 17.02, p = 0.0004]$ as were the main effects of age group $[F(1,24) = 16.99, p = 0.0004]$ and masker $[F(1,24) = 28.82, p < 0.0001]$. *Post hoc* *t*-tests with Bonferroni correction indicated that the effect of masker was significant for both infants $[t(15) = -2.20, p = 0.02]$ and adults $[t(9) = -3.97, p = 0.0016]$. Thus the significant interaction indicates that adults' MUD is greater than that of infants.

Experiment 3 demonstrated two things. First, the apparent decrease in performance in vowel discrimination in the
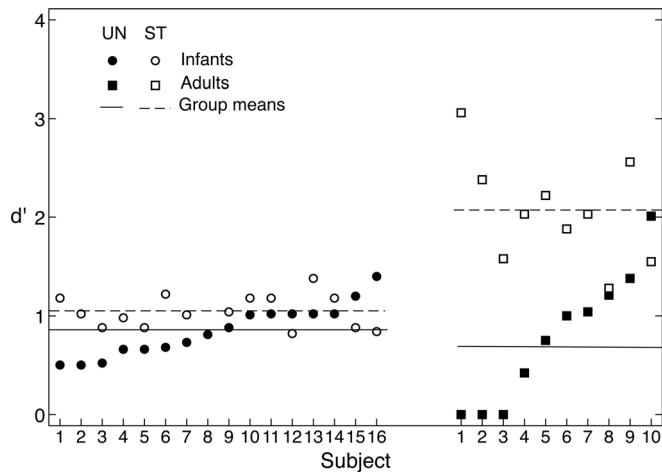
FIG. 3. d' in vowel discrimination for individual subjects for ST modulated and UN noise maskers from experiment 3. Level of the tone was the same for all subjects in an age group; masker condition order counterbalanced across subjects. Solid line represents mean d' in UN noise masker and dashed line represents mean d' in ST modulated noise masker by age group.

ST masker in experiment 1 may well have been due to the procedure of testing the ST masker condition after multiple tests of the UN masker condition. When the order of conditions was counterbalanced across subjects in experiment 3, performance in the ST masker is not worse than that in the UN masker. Second, it appears that infants have a positive, if small, MUD.

V. EXPERIMENT 4: VOWEL DISCRIMINATION IN SAM NOISE

Recall that Newman (2009) reported that infants were worse at recognizing target speech in competition with single-talker speech compared to multitalker speech. She suggested that infants could be distracted from the target by the spectra/temporal variation in single-talker speech. Although ST noise does not have the spectral variation over time associated with speech, its envelope does vary unpredictably over time. The purpose of experiment 4 was to determine if infants' MUD would be greater for SAM noise. If infants are able to listen in the dips but are nonetheless distracted by the speechlike quality of the ST masker, then they may show greater MUD when the masker modulation is not speechlike. The regularity and predictability of SAM noise may also make it easier for infants to take advantage of masker fluctuations. A slow SAM rate with deep modulation was used to ensure that immature temporal resolution did not limit their performance in the modulated masker. A slow modulation rate also ensured that, on average, about half of the duration of each vowel presented would be "unmasked" in the modulated masker condition.

A. Methods

1. Stimuli

The same vowels were used to assess vowel discrimination as in the previous experiments. The level of the vowels was fixed at 56 dB SPL for infants and at 44 dB SPL for adults. The UN masker was the same as in the previous

experiments. The SAM masker was the speech spectrum noise used in the previous experiments, modulated by a raised 4-Hz sinusoid with 75% or 100% modulation depth (m). The RMS amplitude of the SAM masker waveform was matched to that of the unmodulated masker; the level of the maskers was fixed at 60 dB SPL. The starting phase of the modulator varied randomly across sessions. Only infants were tested at 100% m because it was difficult to find a signal level for adults that produced a $d' \sim 1$ in the UN masker that also produced less than perfect performance in the SAM masker at 100% m.

2. Subjects

Data were obtained from 24 infants and 10 adults. Fourteen infants were tested at 75% m and 10 infants were tested at 100% m. The infants' average age was 29.6 wk (SD = 3.0); the adults' average age was 22.1 yr (SD = 1.9). The inclusion criteria were the same as in the previous experiments. Thirteen additional infants did not reach training criteria; 11 of the 13 were tested first in the UN masker condition. Seven infants completed training but did not complete a test condition; six of the seven were tested in the UN masker condition first. Eleven infants completed only one test condition; 7 of the 11 were tested in the UN masker condition first.

3. Procedure

The basic procedure was the same as that in experiment 3: Subjects were tested in both the UN and SAM masker with order counterbalanced across subjects. All subjects had d' greater than 0.8 in the SAM masker condition. The average number of trials to criterion for infants was 22.7 (SD = 9.9) in the SAM masker conditions and 25.5 (SD = 13.4) in the UN masker condition. The average number of trials to criterion for adults was 8.7 (SD = 0.5) in the SAM masker condition and 10.8 (SD = 3.8) in the UN masker condition.

B. Results and discussion

Individual d' is plotted for infants tested at 75% m, infants tested at 100% m, and adults tested at 75% m in Fig. 4 with mean d' indicated by the horizontal lines. Both age groups performed better in the SAM masker than in the UN masker, although the infants' MUD appeared to be smaller than the adults'. Eleven of 14 infants tested at 75% m, 8 of 10 infants tested at 100% m, and all 10 adults had higher d' in the SAM masker condition than in the UN masker condition. An age group × masker ANOVA of d' confirmed the statistical significance of these trends in the 75% m condition: The age group × masker interaction [$F(1,24) = 99.3$, $p < 0.0001$] was statistically significant as were the main effects of age group [$F(1,24) = 76.8$, $p < 0.0001$] and masker [$F(1,24) = 150.8$, $p < 0.0001$]. Post hoc paired t-tests of d' with Bonferroni correction for the individual age groups indicated that the masker effect was significant for both age groups, infants [$t(15) = -4.91$, $p = 0.0002$] and adults [$t(9) = -11.53$, $p < 0.0001$]. Thus the significant interaction indicates that effect of masker condition was greater for adults than for infants.
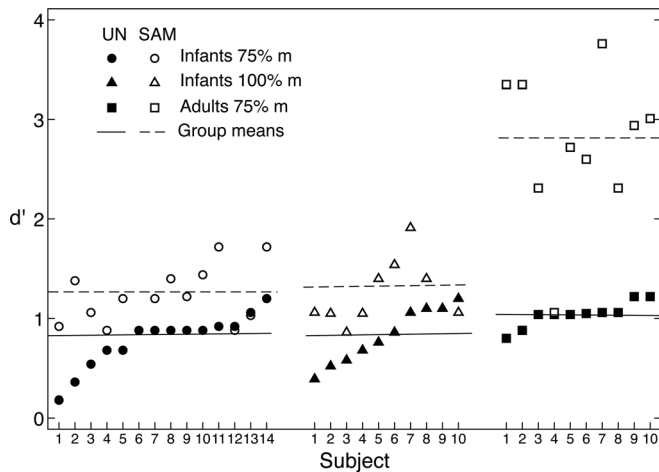
FIG. 4. d' in vowel discrimination for individual subjects for sinusoidally amplitude modulated (SAM) noise maskers at two modulation depths (m) and for UN noise maskers from experiment 4. Level of the tone was the same for all subjects in an age group; masker condition order counterbalanced across subjects. Only infants were tested with 100% modulation depth SAM noise. Solid line represents mean d' in UN noise masker and dashed line represents mean d' in SAM noise masker by age group and modulation depth.

For those infants who completed a test block in only one masker condition, mean d' in the UN masker condition was 0.78 (SD = 0.56) and in the SAM masker condition was 0.65 (SD = 0.19). However, one infant in the UN masker condition had a $d' = 1.9$. Without that subject's data, the mean d' was 0.62 (SD = 0.35). Thus the between-subject comparison, based on a small n, showed no difference between the masker conditions for infants although it also suggests that the infants who finished only one condition were less sensitive to the vowel change at 56 dB SPL than the infants who completed both conditions.

If infants' small MUD is the result of poor temporal resolution, increasing the modulation depth should increase the MUD for infants. As shown in Fig. 4, while infants' MUD was slightly greater for 100% m (0.32) than for 75% m (0.24), the difference was not statistically significant. A masker × modulation depth ANOVA of d' indicated that the masker × modulation depth interaction was not statistically significant [$F(1,25) = 0.01$, $p = 0.9408$]; the main effect of masker [$F(1,25) = 20.52$, $p = 0.0001$] was statistically significant, but the main effect of modulation depth was not statistically significant [$F(1,25) = 0.09$, $p = 0.7682$].

To determine whether infants' relative improvement differed between the ST and SAM masker conditions, a MUD was calculated for each subject in experiments 3 and 4. Because adults were not tested at 100% m, only the 75% m data were included in the analysis. The infants' MUD averaged 0.2 in the ST masker and 0.4 in the SAM masker; the adults' MUD averaged 1.3 in the ST masker and 1.7 in the SAM masker. An age group × modulation type (ST vs SAM) ANOVA indicated a marginally significant effect of modulation type [$F(1,45) = 3.78$, $p = 0.0581$] and a significant effect of age [$F(1,45) = 42.04$, $p < 0.0001$]. The age group × modulation type interaction was not significant [$F(1,45) = 0.07$, $p = 0.7919$]. Thus there was a trend toward a greater MUD in the SAM masker than in the ST masker

for both age groups. Thus the regularity, predictability, and low modulation rate of the SAM noise may have helped listeners to take greater advantage of masker modulation.

The results of this experiment suggest that either increasing the regularity of modulation or decreasing the rate of modulation or both may increase infants' and adults' MUD because there was a trend for MUD to be greater for the SAM masker than for the ST masker. Because increasing the modulation depth of the SAM masker did not increase the MUD, the results also suggest that immature temporal resolution does not limit infants' MUD. However, even with the SAM masker, infants' MUD is smaller than that of adults. One possible explanation is that infants' MUD is limited by a ceiling effect because their best attainable performance is less than perfect. Experiment 5 addressed this possibility.

VI. EXPERIMENT 5: ASYMPTOTIC MASKED VOWEL DISCRIMINATION

The upper asymptote of infants' and children's psychometric functions for detection of a tone in noise is about 85%–90% correct (Bargones and Werner, 1994; Werner and Boike, 2001; Wightman and Allen, 1992). Nozza *et al.* (1990) described infants' psychometric function for consonant discrimination in noise. The upper asymptote of the function was also around 85% correct. It is often assumed that a psychometric function with an upper asymptote less than 1 indicates listener inattentiveness (e.g., Green, 1995). A simple model holds that when the listener is perfectly attentive, there is a sound level at which discrimination approaches perfection. If the listener is not perfectly attentive, on some proportion of trials, he or she obtains no information about the stimulus and guesses as to whether the signal occurred. Thus at that sound level, the probability of a correct response is close to 1 when the listener is attending to the stimulus, and the probability of a correct response of two alternatives is 0.5 when the listener is not attending to the stimulus.[1] If the inattention rate is p_i, then an unbiased listener will achieve an upper asymptote $p(C) = (1 - p_i) + 0.5 p_i$. Although this is not the only possible model, it provides an adequate description of infants' and children's psychometric functions for detection in noise (Bargones *et al.*, 1995; Werner and Boike, 2001; Wightman and Allen, 1992). Furthermore, Nozza *et al.* (1990) report that infants' psychometric function for consonant discrimination in noise is similar to their psychometric function for tone detection in noise.

In the previous experiments, adults achieved $d' \sim 2.7$ in the modulated masker conditions. That translates into $p(C)_{max} \sim 0.9$, where $p(C)_{max}$ is the proportion correct achieved by an unbiased listener with a given d'. If infants obtain as much benefit from modulation as adults and if the best performance they achieve is $p(C)_{max} \sim 0.85$, as suggested by published studies (e.g., Nozza *et al.*, 1990), then they should be able to achieve $d' \sim 2$, in contrast to $d' \sim 1.25$ that they achieved in the modulated masker conditions in experiments 3 and 4. If infants' asymptotic performance in masked vowel discrimination is poorer than that reported in previous studies, then their MUD would be

restricted. The purpose of experiment 5 was to estimate the upper asymptote of infants' psychometric function for masked vowel discrimination.

A. Methods

1. Stimuli

The vowels were the same as those used in the previous experiments. Three groups of infants were tested. One group was tested with the vowels at 75 dB SPL in the unmodulated speech-spectrum noise. The other two groups were tested with 4-Hz SAM speech-spectrum noise with 100% m. One group heard the vowels at 65 dB SPL, the other at 75 dB SPL. The infants tested in modulated noise were tested before those in UN. The 65-dB-SPL infants were tested before the 75-dB-SPL infants.[2] The level of the masker was 60 dB SPL as in the previous experiments.

2. Subjects

Only infant subjects were tested in this experiment, because there is no question that adults' asymptotic $p(C)_{max}$ in this task would be 1. Ten infants completed testing in each group. Their average age was 29.5 wk (SD = 1.9). Five additional infants did not complete training; one of these infants was tested in UN, three with 65-dB vowels in SAM noise, and two in 75-dB vowels in SAM noise. Ten additional infants did not complete all test trials; one of these infants was tested in UN, seven with 65-dB vowels, and two with 75-dB vowels in SAM noise.

3. Procedure

The procedure was generally the same as in the previous experiments. However, the same vowel level was used in the training and test phases of the session, either 65 or 75 dB SPL. Once the training criterion had been met, 20 more trials were completed at the same level, 10 change and 10 no-change trials, randomly ordered. No probe trials were included. The calculation of d' was based on 30 trials: The last 10 trials of training—the trials on which the infant reached the 80% correct criterion—and the 20 test trials. This procedure was meant to help ensure that the infants were at their best on all trials, in case interest waned on later trials, but could bias the results toward higher values of d'. The average number of trials to criterion was 25.8 (SD = 10.4) in the SAM masker conditions and 25.6 (SD = 11.8) in the UN masker condition.

B. Results

Individual d' is plotted for the unmodulated masker at 75 dB SPL and as a function of vowel level for the modulated maskers in Fig. 5, where the short horizontal lines represent mean d' in each condition. The data at 56 dB SPL are those of the 11 infants in experiment 4 who were tested at 100% m, replotted from Fig. 4. Mean d' was about the same at all levels with the modulated masker. A one-way ANOVA of d', including the data from all three vowel levels in the SAM masker conditions, indicated that the effect of level
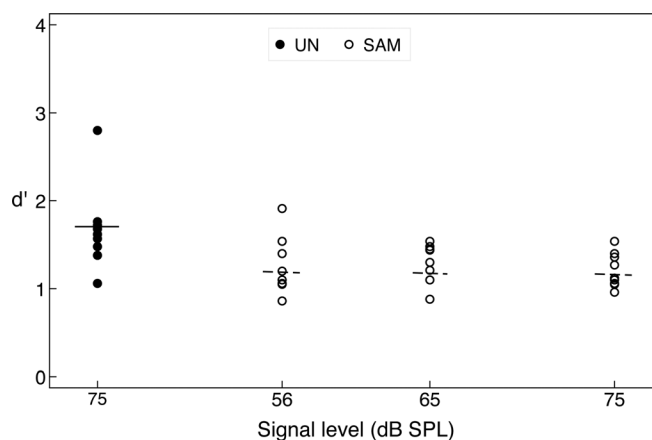

FIG. 5. d' in vowel discrimination by infants as a function of masker type (UN, SAM at 100% depth) and signal level from experiment 5. Solid line represents mean d' in UN noise masker and dashed line represents mean d' in SAM noise masker by condition.

was not statistically significant [$F(2,28) = 0.01$, $p = 0.9915$]. Thus infants' performance does not appear to improve at all at levels above 56 dB SPL.

It appears that d' was higher in the unmodulated masker than in the modulated maskers. One subject achieved $d' \sim 2.8$ in the unmodulated masker. A t-test comparing average d' of infants tested with the unmodulated masker to that of infants tested with the modulated masker at 75 dB SPL confirmed the significance of this difference [$t(18) = 2.68$, $p = 0.0154$]. If the single outlier subject was removed from the unmodulated masker sample, the difference between groups remained significant [$t(17) = 3.07$, $p = 0.0070$]. Thus the upper asymptote of the infants' psychometric function for masked vowel discrimination is lower for a modulated masker than for an unmodulated masker.

C. Discussion

The results of experiment 5 suggest that the upper asymptote of infants' psychometric function for vowel discrimination in modulated noise is lower than that in UN. The upper asymptote of the infants' psychometric function for vowel discrimination in UN was lower than predicted from previous studies, around $0.8 \, p(C)_{max}$ compared to 0.85 reported in tone detection (e.g., Bargones et al., 1995) and in consonant discrimination (Nozza et al., 1990). It is unlikely that the difference between the current and previous studies is due solely to the test method, as previous studies of discrimination and detection using this method have observed asymptotic performance of $0.85 \, p(C)_{max}$ or better (e.g., Bargones et al., 1995; Olsho et al., 1987; Werner and Boike, 2001). Synthetic vowels were used in the current study, but were also used by Nozza et al. (1990) who reported an upper asymptote around $0.85 \, p(C)$.

It is possible that stimulus duration had some effect on infants' asymptotic performance. The vowels used here were 200 ms long, and in the modulated masker, part of the vowel would have often been obscured. In the studies of tone detection and discrimination in which asymptotic performance

J. Acoust. Soc. Am., Vol. 133, No. 6, June 2013

Lynne A. Werner: Modulated-unmodulated difference in infants   4163

was examined, tone durations of 300 (Bargones *et al.*, 1995; Werner and Boike, 2001) to 500 ms (Olsho *et al.*, 1987) were used. Nozza *et al.* (1990) used 300-ms long stimuli in their study of infant consonant discrimination in noise. Bargones *et al.* also reported that infants' psychometric function for detection of repeated 16-ms long tones in quiet was much shallower than that for detection of repeated 300 ms tones, but whether the stimulus duration used here is short enough to account for the difference between the current study and previous studies in asymptotic performance is not clear.

That infants' discrimination in the SAM masker did not improve with increases in target level raises the possibility that infants' ability to listen in masker dips, at least for the 100% SAM masker, is comparable to that of adults. If infants are discriminating vowels at close to asymptotic levels in modulated noise, and if the upper asymptote is reduced as a result of a process like inattentiveness, then on trials on which infants are on task, they are responding nearly perfectly.

VII. GENERAL DISCUSSION

These studies demonstrate that a positive MUD can be observed in infants at least under some conditions. Infants' MUD was found to be smaller than that of adults given approximately the same baseline levels of performance in the unmodulated masker condition. Infants' high-level vowel discrimination in modulated noise was poorer than that in UN; that is, the upper asymptote of infants' psychometric function for masked vowel discrimination appears to be lower for a SAM masker than for an unmodulated masker. Thus while infants' masked vowel discrimination improves in modulated relative to UN, the extent of improvement appears to be limited by a low performance ceiling.

In the current study, infants achieved a $d' \sim 1$ in masked vowel discrimination at a target-to-masker ratio (TMR) of about -4 dB in UN. That value is close to that reported for masked consonant discrimination by Nozza *et al.* (1990), about -3 dB, for infants of comparable age. Similarly, 7-month-old infants here achieved a $d' \sim 1$ for masked detection of a 50-ms tone at the same TMR as reported for infants of the same age by Bargones *et al.* (1995) for a 16-ms tone. Note that adults' performance in the current experiment is also about the same as reported in those two studies of infants. Thus both infants' and adults' masked discrimination and detection in UN here are as expected on the basis of previous work.

Presumably, infants in the current study could have achieved a $d' \sim 1$ at TMR lower than -4 dB in the ST and SAM noise maskers because they were more sensitive to the vowel change in the modulated than in the unmodulated masker. Truly comparable data for infants' discrimination and detection in a modulated masker are not available; studies examining infants' speech recognition in a background of single-talker speech come closest in that regard. Newman (2009), discussed in the preceding text, found that 5 and 8.5-month-old infants did not recognize their names, produced by female talkers in a background of fluent speech produced by a different single female talker at 10-dB TMR. Newman and Juscyk (1996) familiarized 7.5-month-old infants with a series of words spoken by a female talker in quiet. Infants' recognition of those words was then tested with the words presented in a background of fluent speech produced by a single male talker. Infants' recognized the familiar words at 5 dB, but not 0 dB, TMR. If visual information, synchronized with the target words, was also available, infants succeeded in the same task at 0 dB TMR (Hollich *et al.*, 2005). Although these studies used a different methodology from that employed in the current study, their results suggest that under favorable circumstances, infants can recognize speech at 0-dB TMR in a single-talker masker. Although lower TMRs were not tested, the results of these studies are not inconsistent with the results of the current study.

The upper asymptote of the psychometric function for vowel discrimination observed here was lower than those previously reported for other tasks (e.g., Bargones *et al.*, 1995; Nozza *et al.*, 1990). Infants who completed training and testing in experiment 5, overall, achieved only $0.73 \, p(C)_{max}$ on average in modulated noise and $0.80 \, p(C)_{max}$ on average in UN. The latter finding raises the possibility that the 80% correct training criterion used in all of these experiments is overly restrictive, eliminating infants with lower levels of asymptotic performance from the final data set and overestimating the infants' average discrimination capacity in all of the conditions tested. Because asymptotic performance was poorer in modulated than in UN, in fact, it is possible that only some infants have a positive MUD and that infants with little or no MUD were excluded from the final dataset. That possibility cannot be eliminated; however, if the 80% correct training criterion eliminated less sensitive infants, the proportion of infants not reaching criterion should be greater than that in other tasks with higher upper asymptote but the same training criterion. In fact, the percentage of infants not reaching training criterion—24%, 22%, and 11% in experiments 3, 4, and 5, respectively—is not atypical in studies of this type. In recent observer-based studies of tone-in-noise detection, for example, 15%–33% of infant subjects did not reach the 80% correct training criterion (Werner and Boike, 2001; Werner *et al.*, 2009). Asymptotic performance in that task is about $0.85 \, p(C)_{max}$. Moreover, even though the upper asymptote of the psychometric function for vowel discrimination in modulated noise was significantly lower than that in UN in the current study, the proportion of infants reaching training criterion was greater in the modulated noise condition than in the UN condition. Similarly, the average number of trials required for infants to reach training criteria in the current experiments, ranging from 22 to 28 trials, is comparable to that reported in previous studies (20.3 trials, Werner and Boike, 2001; 22 trials, Werner *et al.*, 2009). Thus there is no clear evidence that the training criterion used in these experiments eliminated an atypical number of infants from the final dataset.

The major question raised by the results of the current study is why infants do not perform better in vowel discrimination in a modulated masker. In the introduction, several factors that might limit infants' ability to take advantage of

masker modulation were described. Each of these factors is considered in the following paragraphs.

A. Immature temporal resolution

Although the modulation rate used here was quite low, the current finding that increasing the depth of a SAM masker did not improve infant's vowel discrimination argues against the importance of immature temporal resolution in restricting infants' MUD. That conclusion is consistent with those of studies reporting that infants' temporal resolution appears mature by some measures (Werner, 1999, 2006) and that young children, with mature TMTF shape, still obtain a smaller advantage of masker temporal modulation than adults (Grose *et al.*, 1993; Hall *et al.*, 2012; Hall and Grose, 1994).

B. Higher target-to-masker ratio required by infants, compared to adults, in unmodulated masker

At higher TMR, the advantage of masker modulation is reduced for adults (e.g., Bernstein and Grant, 2009; Oxenham and Simonson, 2009). For example, Bernstein and Grant (2009) found that when tested at the TMR that yielded 50% correct word recognition in UN for each listener, hearing-impaired adults had smaller average MUD in sentence recognition than normal hearing adults. However, when tested at the TMR required by the hearing-impaired adults to reach 50% correct in UN, normal hearing listeners had nearly as small a MUD as the hearing-impaired listeners. Hall *et al.* (2012) provided some evidence that the higher TMR at threshold in UN could account for younger children's reduced MUD in sentence recognition. One explanation for the effect of TMR on the MUD in sentence recognition is that speech cues are distributed across a range of levels and that at high levels, more cues are masked at the peaks in the masker relative to the number of cues unmasked in the dips of the modulated masker. It has also been suggested that when the modulated masker dominates the auditory response (TMR > 0 dB), masker fluctuations can distort the target waveform, partially offsetting the advantage of intermittent improvements in TMR (Oxenham and Simonson, 2009).

When the task is discrimination between two vowels, however, it is not clear that the level-distribution-of-speech-cues explanation for reduced MUD at high TMR applies. It is also not clear that the modulated masker would dominate the auditory response at −4 dB TMR, the TMR used for infants here. A ceiling effect prevents the assessment of adults' MUD at −4 dB TMR: Adults' performance in both the unmodulated and modulated masker conditions is perfect (unpublished observations). That adults do not perform more poorly in the modulated masker than in the unmodulated masker at −4 dB argues against the idea that masker fluctuations substantially distort the target waveform at that TMR.

C. Inefficient dip listening

Infants could have some trouble identifying the time points in the modulated masker when the TMR is advantageous.

Grose *et al.* (1993) hypothesized that such inefficient dip listening could account for the fact that children had a smaller MUD than adults in tone detection. A finding of the current study that is consistent with a possible role of inefficient dip listening is the trend—just missing statistical significance—for listeners to obtain a greater MUD for the more predictable SAM than for the ST masker. However, infants' and adults' MUD increased equivalently with the change in modulation type, whereas if infants were relatively inefficient dip listeners, it would be expected that they would benefit more from increasing the regularity. Furthermore, even the regularity of SAM was insufficient to raise infants' MUD to the value achieved by adults. Finally, one would think that raising the TMR by nearly 20 dB would be sufficient to overcome the effects of inefficient dip listening, when in fact, infant performance was not better for 75 dB SPL than for 56 dB SPL targets.

D. Immature ability to use partial information

Infants' small MUD relative to adults' may result from adults' superior ability to recognize a target on the basis of "partial information" (Hall *et al.*, 2012; Stuart, 2005; Stuart *et al.*, 2006; but see Johnson, 2004). This idea is typically applied to tasks in which the listener recognizes words or sentences in a modulated masker: To understand a word that is intermittently masked by a modulated masker, the listener must integrate acoustic cues and other information across glimpses of the target that are audible in the low-level portions of the masker. When the target is a brief tone or a steady-state vowel, in which there is no change in the target over time, the issue is probably less about combining information across glimpses than about identifying a short duration target, particularly when the duration of the target is short compared to the period of the modulator. There is an 80% chance that some portion of the 200-ms vowel falls at a minimum in a 4-Hz modulated masker.

The threshold of 7-month-olds for detecting a tone in quiet is higher for a 16-ms long tone than for a 300-ms tone, by nearly 30 dB, far more than seen in adults (Bargones *et al.*, 1995), suggesting that infants have difficulty detecting very short duration sounds. It may be that infants' ability to make use of information in masker dips is as good as that of adults but that the reduced target duration makes it difficult for them to identify the vowel and thus offsets the advantage of dip listening. Whether the reduction in the audible duration imposed by the modulated masker here is sufficient to offset the benefit of modulation for infants is not known. Adults' vowel identification is apparently based on the output of a short (<50 ms) temporal window (Wallace and Blumstein, 2009), but the effect of duration on infants' vowel discrimination has not been examined. However, if infants were differentially sensitive to a reduction of vowel duration, their performance in the single-talker masker condition would be expected to be poorer than that in the SAM masker because the single-talker masker contains higher frequency modulations and therefore shorter glimpses of the vowels. The difference between those masker conditions was marginally significant here, but there was no evidence that the difference was greater for infants than for adults.

E. Informational masking or distraction

When the competing sound is single-talker speech, the spectral-temporal variation in the masker is likely to cause informational masking (Brungart and Simpson, 2007). For a speech target in single-talker speech background, the similarity of the target and background may also promote informational masking. Infants are known to be more susceptible than adults to informational masking even in the absence of masker variation (Leibold and Werner, 2006). Newman (2009) suggested that informational masking could account for infants' poor recognition of speech in a single-talker speech background. That explanation of Newman's results is bolstered by the fact that introducing target-masker differences in fundamental frequency (Newman and Jusczyk, 1996) or adding audiovisual cues to the target (Hollich et al., 2005) apparently improved infants' speech recognition. Manipulations such as these are known to reduce informational masking more than they reduce energetic masking (Darwin et al., 2003; Wightman et al., 2006). The results of the current study could indicate that temporal variation alone is sufficient to induce informational masking in infants despite the spectral shape and temporal property differences between the target vowel and the modulated masker. The slope of the psychometric function in informational masking is known to be very shallow (Kidd et al., 2003; Lutfi et al., 2003; Neff and Callaghan, 1988). What appears to be a reduced upper asymptote of the psychometric function for masked vowel discrimination in a modulated masker by infants may actually a reflect a very, very shallow psychometric function.

Brungart (2001) argued that speech-on-speech masking in the coordinate response measure paradigm is dominated by informational masking on the basis of the observation that subjects' errors are most frequently reporting what the masker talker said rather than what the target talker said. In other words, listeners responded to the masker rather than to the target. A related explanation of infants' difficulties separating speech from a single-talker masker, also proposed by Newman (2009), was distraction, the idea that infants are responding to the masker rather than the target on some trials, presumably because the modulated masker holds some inherent interest for them. Infants' (marginally) smaller MUD in the ST masker condition than in the SAM masker condition here would be consistent with the idea that the ST masker is more distracting to the infants than the SAM masker. However, why the same reduction in MUD for the ST masker relative to the SAM masker would be seen in adults is less clear.

Distraction might be modeled as what Wightman and Allen (1992) describe as forgetting: On some proportion of trials, the listener "forgets" which sound is the target and responds instead to the masker. Forgetting in the current context would mean that infants are responding to the modulated masker on some proportion of trials rather than to the target vowel. The upper asymptote of the forgetful listener's psychometric function is equal to his forgetting rate. That the upper asymptote of the infants' psychometric function for vowel discrimination in a modulated masker is lower than that in an

unmodulated masker is consistent with the idea that infants respond more to (or are more distracted by) the modulated masker than the unmodulated masker. One experiment that would test this hypothesis would be one in which the relative attractiveness of the masker is manipulated. Of course, a combination of inattentiveness and distraction may finally be required to account for the upper asymptote of infants' psychometric function for vowel discrimination in modulated noise.

VIII. CONCLUSIONS

Infants discriminate between vowels better in a modulated masker than in an unmodulated masker as long as the order of testing is controlled. The difference between infants' sensitivity to a vowel change in modulated and unmodulated maskers is smaller than the difference observed for adults. Infants' vowel discrimination may be better in a slow, deep SAM masker than in a ST masker, suggesting that more regular or slower modulations improve their ability to use information in the dips of the modulated masker. However, the improvement is no greater for infants than for adults and thus cannot explain infants' relatively small MUD. Increasing the target-to-masker ratio has little effect on infants' sensitivity to a vowel change in a SAM noise masker, suggesting that infants' MUD is limited by a low performance ceiling. Informational masking or distraction, combined with inattentiveness, may explain the low upper asymptote of infants' psychometric function for vowel discrimination, and hence, their small MUD.

[1]In the observer-based procedure, it is not possible to separate the contributions of the observer and the infant listener to the outcome on each trial. However, the psychometric functions of of infants tested in the observer-based procedure are very similar to those of infants in a conditioned head-turn procedure (Nozza et al., 1990) and of children tested in a three-alternative forced-choice procedure (Wightman and Allen, 1992). Because the listener's response in the latter two cases does not depend on an observer, we believe that observer-based psychometric functions reflect the characteristics of the infant listener.

[2]The expectation was that infants' performance would be asymptotic with 65 dB SPL vowels because we had generally been successful at training infants to an 80% correct criterion at that level. When the average performance of infants at 65 dB SPL on 30 trials fell short of 80% correct, a second group of infants was tested at 75 dB SPL. When the average performance of the infants tested at 75 dB SPL still fell short of 80% correct, a third group of infants was test with the unmodulated masker with 75 dB SPL vowels.

Bacon, S. P., and Lee, J. (**1997**). "The modulated-unmodulated difference: Effects of signal frequency and masker modulation depth," J. Acoust. Soc. Am. **101**, 3617–3624.

Bacon, S. P., Lee, J., Peterson, D. N., and Rainey, D. (**1997**). "Masking by modulated and unmodulated noise: Effects of bandwidth, modulation rate, signal frequency, and masker level," J. Acoust. Soc. Am. **101**, 1600–1610.

Bargones, J. Y., and Werner, L. A. (**1994**). "Adults listen selectively; infants do not," Psychol. Sci. **5**, 170–174.

Bargones, J. Y., Werner, L. A., and Marean, G. C. (**1995**). "Infant psychometric functions for detection: Mechanisms of immature sensitivity," J. Acoust. Soc. Am. **98**, 99–111.

Bernstein, J. G. W., and Grant, K. W. (**2009**). "Auditory and auditory-visual intelligibility of speech in fluctuating maskers for normal-hearing and hearing-impaired listeners," J. Acoust. Soc. Am. **125**, 3358–3372.

Bor, S., Souza, P., and Wright, R. (**2008**). "Multichannel compression: Effects of reduced spectral contrast on vowel identification," J. Speech Lang. Hear. Res. **51**, 1315–1327.

Brungart, D. S. (**2001**). "Informational and energetic masking effects in the perception of two simultaneous talkers," J. Acoust. Soc. Am. **109**, 1101–1109.

Brungart, D. S., and Simpson, B. D. (**2007**). "Effect of target-masker similarity on across-ear interference in a dichotic cocktail-party listening task," J. Acoust. Soc. Am. **122**, 1724–1734.

Buss, E., Wall, J. W., and Grose, J. H. (**2003**). "Effect of amplitude modulation coherence for masked speech signals filtered into narrow bands," J. Acoust. Soc. Am. **113**, 462–467.

Cox, R. M., Alexander, G. C., and Gilmore, C. (**1987**). "Development of the Connected Speech Test (CST)," Ear Hear. **8**, S119–S126.

Darwin, C. J., Brungart, D. S., and Simpson, B. D. (**2003**). "Effects of fundamental frequency and vocal-tract length changes on attention to one of two simultaneous talkers," J. Acoust. Soc. Am. **114**, 2913–2922.

Drullman, R., and Bronkhorst, A. W. (**2004**). "Speech perception and talker segregation: Effects of level, pitch, and tactile support with multiple simultaneous talkers," J. Acoust. Soc. Am. **116**, 3090–3098.

Dubno, J. R., Horwitz, A. R., and Ahlstrom, J. B. (**2002**). "Benefit of modulated maskers for speech recognition by younger and older adults with normal hearing," J. Acoust. Soc. Am. **111**, 2897–2907.

Eisenberg, L. S., Dirks, D. D., and Bell, T. S. (**1995**). "Speech recognition in amplitude-modulated noise of listeners with normal and listeners with impaired hearing," J. Speech Lang. Hear. Res **38**, 222–233.

Fernald, A., Swingley, D., and Pinto, J. P. (**2001**). "When half a word is enough: Infants can recognize spoken words using partial phonetic information," Child Dev. **72**, 1003–1015.

Festen, 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. (**2004**). "Effect of number of masking talkers and auditory priming on informational masking in speech recognition," J. Acoust. Soc. Am. **115**, 2246–2256.

Green, D. M. (**1995**). "Maximum likelihood procedures and the inattentive listener," J. Acoust. Soc. Am. **97**, 3749–3760.

Grose, J. H., Hall, J. W., and Gibbs, C. (**1993**). "Temporal analysis in children," J. Speech Lang. Hear. Res. **36**, 351–356.

Hall, J. W., Buss, E., Grose, J. H., and Roush, P. A. (**2012**). "Effects of age and hearing impairment on the ability to benefit from temporal and spectral modulation," Ear Hear. **33**, 340–348.

Hall, J. W., 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.

Hollich, G., Newman, R. S., and Jusczyk, P. W. (**2005**). "Infants' use of synchronized visual information to separate streams of speech," Child Dev. **76**, 598–613.

Howard-Jones, P. A., and Rosen, S. (**1993**). "The perception of speech in fluctuating noise," Acustica **78**, 258–272.

Johnson, S. P. (**2004**). "Development of perceptual completion in infancy," Psychol. Sci. **15**, 769–775.

Kidd, G., Mason, C. R., and Richards, V. M. (**2003**). "Multiple bursts, multiple looks, and stream coherence in the release from informational masking," J. Acoust. Soc. Am. **114**, 2835–2845.

Kwon, B. J., and Turner, C. W. (**2001**). "Consonant identification under maskers with sinusoidal modulation: Masking release or modulation interference?" J. Acoust. Soc. Am. **110**, 1130–1140.

Leibold, L. J., and Werner, L. A. (**2006**). "Effect of masker-frequency variability on the detection performance of infants and adults," J. Acoust. Soc. Am. **119**, 3960–3970.

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.

Neff, D. L., and Callaghan, B. P. (**1988**). "Effective properties of multicomponent simultaneous maskers under conditions of uncertainty," J. Acoust. Soc. Am. **83**, 1833–1838.

Nelson, P. B., Jin, S. H., Carney, A. E., and Nelson, D. A. (**2003**). "Understanding speech in modulated interference: Cochlear implant users and normal-hearing listeners," J. Acoust. Soc. Am. **113**, 961–968.

Newman, R. S. (**2009**). "Infants' listening in multitalker environments: Effect of the number of background talkers," Atten. Percept. Psychophys. **71**, 822–836.

Newman, R. S., and Jusczyk, P. W. (**1996**). "The cocktail party effect in infants," Atten. Percept. Psychophys. **58**, 1145–1156.

Nozza, R. J., Rossman, R. N. F., Bond, L. C., and Miller, S. L. (**1990**). "Infant speech-sound discrimination in noise," J. Acoust. Soc. Am. **87**, 339–350.

Olsho, L. W., Koch, E. G., and Halpin, C. F. (**1987**). "Level and age effects in infant frequency discrimination," J. Acoust. Soc. Am. **82**, 454–464.

Oxenham, A. J., and Simonson, A. M. (**2009**). "Masking release for low-and high-pass-filtered speech in the presence of noise and single-talker interference," J. Acoust. Soc. Am. **125**, 457–468.

Peters, R. W., Moore, B. C. J., and Baer, T. (**1998**). "Speech reception thresholds in noise with and without spectral and temporal dips for hearing-impaired and normally hearing people," J. Acoust. Soc. Am. **103**, 577–587.

Stuart, A. (**2005**). "Development of auditory temporal resolution in school-age children revealed by word recognition in continuous and interrupted noise," Ear Hear. **26**, 89–95.

Stuart, A. (**2008**). "Reception thresholds for sentences in quiet, continuous noise, and interrupted noise in school-age children," J. Am. Acad. Audiol. **19**, 135–146.

Stuart, A., Givens, G. D., Walker, L. J., and Elangovan, S. (**2006**). "Auditory temporal resolution in normal-hearing preschool children revealed by word recognition in continuous and interrupted noise," J. Acoust. Soc. Am. **119**, 1946–1949.

Summers, V., and Molis, M. R. (**2004**). "Speech recognition in fluctuating and continuous maskers: Effects of hearing loss and presentation level," J. Speech Lang. Hear. Res **47**, 245–256.

Takahashi, G. A., and Bacon, S. P. (**1992**). "Modulation detection, modulation masking, and speech understanding in noise in the elderly," J. Speech Lang. Hear. Res **35**, 1410–1421.

Trehub, S. E., Schneider, B. A., and Henderson, J. (**1995**). "Gap detection in infants, children, and adults," J. Acoust. Soc. Am. **98**, 2532–2541.

Wallace, A. B., and Blumstein, S. E. (**2009**). "Temporal integration in vowel perception," J. Acoust. Soc. Am. **125**, 1704–1710.

Werner, L. A. (**1995**). "Observer-based approaches to human infant psychoacoustics," in *Methods in Comparative Psychoacoustics*, edited by G. M. Klump, R. J. Dooling, R. R. Fay, and W. C. Stebbins (Birkhauser Verlag, Boston), pp. 135–146.

Werner, L. A. (**1999**). "Forward masking among infant and adult listeners," J. Acoust. Soc. Am. **105**, 2445–2453.

Werner, L. A. (**2006**). "Preliminary observations on the temporal modulation transfer functions of infants and adults," paper presented at the American Auditory Society, Scottsdale, AZ.

Werner, L. A., and Boike, K. (**2001**). "Infants' sensitivity to broadband noise," J. Acoust. Soc. Am. **109**, 2101–2111.

Werner, L. A., Marean, G. C., Halpin, C. F., Spetner, N. B., and Gillenwater, J. M. (**1992**). "Infant auditory temporal acuity: Gap detection," Child Dev. **63**, 260–272.

Werner, L. A., Parrish, H. K., and Holmer, N. M. (**2009**). "Effects of temporal uncertainty and temporal expectancy on infants' auditory sensitivity," J. Acoust. Soc. Am. **125**, 1040–1049.

Wightman, F., and Allen, P. (**1992**). "Individual differences in auditory capability among preschool children," in *Developmental Psychoacoustics*, edited by L. A. Werner and E. W. Rubel (American Psychological Association, Washington, DC), pp. 113–133.

Wightman, F., Kistler, D., and Brungart, D. (**2006**). "Informational masking of speech in children: Auditory-visual integration," J. Acoust. Soc. Am. **119**, 3940–3949.