HAL open science

Amplitude modulation detection and modulation masking in school-age children and adults

Emily Buss, Christian Lorenzi, Laurianne Cabrera, Lori J Leibold, John H Grose

To cite this version:

Emily Buss, Christian Lorenzi, Laurianne Cabrera, Lori J Leibold, John H Grose. Amplitude modulation detection and modulation masking in school-age children and adults. Journal of the Acoustical Society of America, 2019, 145 (4), pp.2565-2575. 10.1121/1.5098950 . hal-03898601

HAL Id: hal-03898601

https://hal.science/hal-03898601v1

Submitted on 14 Dec 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Amplitude modulation detection and modulation masking in school-age children and adults

Emily Buss, Christian Lorenzi, Laurianne Cabrera, Lori J. Leibold, and John H. Grose

Citation: The Journal of the Acoustical Society of America **145**, 2565 (2019); doi: 10.1121/1.5098950
View online: https://doi.org/10.1121/1.5098950
View Table of Contents: https://asa.scitation.org/toc/jas/145/4
Published by the Acoustical Society of America

Amplitude modulation detection and modulation masking in school-age children and adults

Emily Buss,[1,a)] Christian Lorenzi,[2] Laurianne Cabrera,[3] Lori J. Leibold,[4] and John H. Grose[1]

[1]*Department of Otolaryngology/Head and Neck Surgery, School of Medicine, University of North Carolina, Chapel Hill, North Carolina 27599-7070, USA*

[2]*Laboratoire des Systèmes Perceptifs, Département d'Études Cognitives, Ecole Normale Supérieure, Universite Paris Sciences et Lettres, Centre National de la Recherche Scientifique, Paris, France*

[3]*Laboratoire de Psychologie de la Perception, Université Paris Descartes, Centre National de la Recherche Scientifique, Paris, France*

[4]*Center for Hearing Research, Boys Town National Research Hospital, Omaha, Nebraska 68131, USA*

(Received 10 January 2019; revised 2 April 2019; accepted 3 April 2019; published online 30 April 2019)

Two experiments were performed to better understand on- and off-frequency modulation masking in normal-hearing school-age children and adults. Experiment 1 estimated thresholds for detecting 16-, 64- or 256-Hz sinusoidal amplitude modulation (AM) imposed on a 4300-Hz pure tone. Thresholds tended to improve with age, with larger developmental effects for 64- and 256-Hz AM than 16-Hz AM. Detection of 16-Hz AM was also measured with a 1000-Hz off-frequency masker tone carrying 16-Hz AM. Off-frequency modulation masking was larger for younger than older children and adults when the masker was gated with the target, but not when the masker was continuous. Experiment 2 measured detection of 16- or 64-Hz sinusoidal AM carried on a bandpass noise with and without additional on-frequency masker AM. Children and adults demonstrated modulation masking with similar tuning to modulation rate. Rate-dependent age effects for AM detection on a pure-tone carrier are consistent with maturation of temporal resolution, an effect that may be obscured by modulation masking for noise carriers. Children were more susceptible than adults to off-frequency modulation masking for gated stimuli, consistent with maturation in the ability to listen selectively in frequency, but the children were not more susceptible to on-frequency modulation masking than adults. © *2019 Acoustical Society of America*. https://doi.org/10.1121/1.5098950

[MD] Pages: 2565–2575

I. INTRODUCTION

Amplitude modulation (AM) is an important source of auditory information. The coherence of AM across frequency influences auditory object formation (Grimault *et al.*, 2002; Nie and Nelson, 2015), and envelope cues support speech perception under some conditions (Drullman *et al.*, 1994a,b; Shannon *et al.*, 1995). Sensitivity to AM can be reduced with the introduction of a competing pattern of AM, whether the target and competing AM patterns are present in the same carrier frequency region (Bacon and Grantham, 1989; Houtgast, 1989) or different frequency regions (Yost and Sheft, 1989; Yost *et al.*, 1989). This general phenomenon is described as modulation masking (Bacon and Grantham, 1989; Sek *et al.*, 2015), although the special case of off-frequency modulation masking is often referred to as modulation detection interference (MDI). Modulation masking has been argued to play an important role in noise-masked speech recognition. By this view, the random inherent envelope fluctuation of nominally steady noise interferes with the perceptual processing of speech envelope cues (Apoux and Bacon, 2008; Stone *et al.*, 2011; Stone *et al.*, 2012; Stone and Moore, 2014). Although modulation masking has been well studied in adults, very little is known about this type of masking in children. The present study was designed to evaluate the development of AM detection with and without a modulation masker.

The pattern of AM detection thresholds as a function of modulation rate is referred to as a temporal modulation transfer function (TMTF). For adults, the TMTF for a wideband noise carrier is bandpass in shape. Thresholds for low-rate AM are limited by the number of modulation cycles present in a listening interval (Sheft and Yost, 1990). Thresholds for high-rate AM fall by approximately 3 dB per octave between 50 and 800 Hz, a result that was originally interpreted as reflecting the limits of temporal resolution (Viemeister, 1979). The TMTF measured with a noise carrier is similar for school-age children and adults, with the caveat that children are less sensitive to AM overall (Hall and Grose, 1994; Park *et al.*, 2015). For example, Hall and Grose (1994) measured sinusoidal AM detection at 5, 20, 100, 150, and 200 Hz using a continuous noise carrier. The similarity in TMTF shape across age groups was interpreted as reflecting similar temporal resolution, and differences in overall sensitivity were interpreted in terms of listening efficiency, defined here as the ability to use the sensory information associated with AM.

Although early reports of TMTFs for noise carriers interpreted the reduced sensitivity to AM with increasing AM rate above 50 Hz as evidence of limited temporal

a)Electronic mail: ebuss@med.unc.edu

resolution, subsequent work has attributed this pattern to modulation masking associated with inherent modulation of the noise carrier (Dau *et al.*, 1997a,b; Dau *et al.*, 1999). For adults, the TMTF obtained using a high-frequency pure-tone carrier differs markedly from that obtained using a noise carrier. For example, sensitivity to AM imposed on a tonal carrier above 2000 Hz is relatively constant for AM rates up to 100–130 Hz and falls above that rate, with threshold improvements observed when the sidebands associated with AM become resolved (Kohlrausch *et al.*, 2000; Stellmack *et al.*, 2005). The observation that the noise-carrier TMTF begins to roll off at a lower rate than the tone-carrier TMTF has been attributed to the detrimental effects of inherent noise fluctuation on the detection of sinusoidal AM, otherwise known as modulation masking. This raises the possibility that the modulation masking associated with noise carriers may foil attempts to characterize the development of temporal resolution using the noise-carrier TMTF paradigm.

Modulation masking is often characterized using a bank of bandpass filters tuned to AM rate, and this approach is successful at characterizing masking effects observed for both within- and across-frequency modulation masking in adults (Verhey and Dau, 2000; Dau *et al.*, 1997a). This model captures the observation that modulation masking is tuned to rate, with the largest modulation masking observed when the target and masker AM are similar in rate (Yost *et al.*, 1989; Ewert and Dau, 2000). When the target and masker carriers are separated in frequency, asynchronous gating improves detection of target AM (Hall and Grose, 1991; Moore and Shailer, 1992; Oxenham and Dau, 2001). A beneficial effect of asynchronous gating of the carriers implies that AM is represented separately in different carrier frequency regions, and off-frequency modulation masking is related to auditory grouping.

Interest in the development of modulation masking is motivated in part by the suggestion that modulation masking limits speech recognition in a nominally steady noise masker (Stone *et al.*, 2011; Stone *et al.*, 2012; Stone and Moore, 2014). Noise-masked speech recognition tends to be immature in school-age children. Some data indicate that children's performance is mature by approximately 10–12 years of age (Stuart, 2008; Corbin *et al.*, 2016), although other data indicate continued development into adolescence (Rashid *et al.*, 2016). Factors proposed to account for the development of speech-in-noise recognition include selective attention, cognitive abilities, and linguistic knowledge (e.g., McCreery *et al.*, 2017), but there is still no consensus regarding the relative contributions of these factors. Greater susceptibility to modulation masking in children was recently proposed as a possible contributor to the development of speech-in-noise recognition (Buss *et al.*, 2018). This would occur if the inherent modulation of a noise masker had a larger detrimental effect on children than adults. This hypothesis is speculative, however, given the lack of published data on modulation masking in children and the preliminary data indicating similar susceptibility to modulation masking in children and adults (Cabrera *et al.*, 2017).

The present experiments sought to clarify the role of modulation masking in children's detection of AM. The first experiment measured AM detection for a pure-tone carrier in children and adults with and without an off-frequency masker tone that was amplitude modulated. The off-frequency masker, when present, was either continuous or gated synchronously with the target. The goal was to characterize both the development of sensitivity to AM as a function of rate in the absence of on-frequency modulation masking and the development of off-frequency modulation masking. The second experiment measured the detection of AM carried on a band of noise with and without additional masker AM patterns applied to that carrier. The goal was to evaluate development of on-frequency modulation masking and tuning to modulation rate. Poorer performance overall was anticipated for children compared to adults, a result that is sometimes described in terms of reduced efficiency (e.g., Hall and Grose, 1994). Greater susceptibility to modulation masking in young children was expected based on development of the ability to listen selectively, which is observed across a wide range of psychoacoustic paradigms (e.g., Leibold, 2012; Leibold and Buss, 2016).

II. GENERAL METHODS

Listeners were normal-hearing 5- to 11-year-olds and adults, with pure-tone thresholds of 20 dB hearing level (HL) or lower between 250 and 8000 Hz bilaterally. Potential participants were excluded if they or their parents reported a history of hearing problems, known developmental or cognitive delays, or a diagnosis of ear infection within the past six months. Listeners were recruited using flyers posted at the university and in the local community. The two experiments described below used different listeners.

In both experiments, the task was to detect AM in a three-alternative forced choice, and thresholds were evaluated using a two-down, one-up stepping rule that estimates the signal strength associated with 71% correct detection. The proportional AM depth (m, 0–1) was adjusted in steps of $20\log(m)$. The maximum target modulation depth was 100% ($m = 1$). For conditions incorporating on-frequency masker AM in experiment 2, the relative AM depth of the target and masker was adjusted while keeping the overall AM depth constant. The step size for adjusting the target AM depth was either 4 or 8 dB at the outset of each track.[1] This was reduced to 2 dB after the second track reversal. Eight reversals were obtained in each track, and the threshold was estimated as the mean AM depth at the last six track reversals. Results reported below are the mean of thresholds from three to four tracks per condition. Within each experiment, data collection was blocked by condition, and conditions were completed in quasi-random order.

The psychophysical test procedures were designed for use with young children. Listening intervals were marked visually by a computer animation, showing three frogs. Each frog opened and closed its mouth synchronously with the associated auditory stimulus. At the end of each trial the listener selected the frog associated with the target AM, using a touchscreen interface. Following the listener's response, feedback was provided wherein the frog associated with the target interval performed a brief animation, consisting of catching and eating a fly. Prior to the first data collection track listeners heard a description of the task, interleaved with examples of the auditory stimuli and video prompts.

This sequence included two example trials with a verbal description of the visual feedback. An abbreviated introduction was presented when stimulus conditions changed, including an example trial with feedback, to orient the listener to the new stimulus condition. This protocol ensured consistency in the instructions provided to each listener and familiarized listeners with the forced-choice task. Listeners were offered frequent breaks, encouragement, and the opportunity to ask questions. Children were rewarded with small toys, and all listeners received $15 per hour. These procedures were approved by the Biomedical Institutional Review Board at the University of North Carolina at Chapel Hill.

Experiments were controlled using custom MATLAB scripts (The MathWorks, Natick, MA). Stimuli were generated in software, played out of a real-time processor (RP2, Tucker-Davis Technologies, Alachua, FL), routed through a headphone buffer (HB7, Tucker-Davis Technologies), and presented diotically over headphones (experiment 1, Sennheiser, HD 25, Wedemark, Germany; experiment 2, Sennheiser, HD 265).

Analyses of age effects were carried out using Pearson correlations, linear mixed models, and *t*-tests. Age was treated as a continuous variable when analyzing data from children; the logarithm of age was used due to the expectation of decelerating effects of development with increasing age. Age was treated as a categorical variable when comparing data from children and adults. For the linear mixed model analyses, child age was centered at 7.5 years, the approximate geometric mean age of the child samples in both experiments, and listener was included as a random factor. A significance criterion of $\alpha = 0.05$ was adopted, and all tests were two-tailed unless otherwise specified. Listener sex was evaluated as a factor in all comparisons reported below and found to be non-significant in all cases.

III. EXPERIMENT 1: AM DETECTION FOR A PURE-TONE CARRIER WITH AND WITHOUT AN OFF-FREQUENCY MODULATION MASKER

A. Methods

Listeners were 25 children between 5.3 and 11.1 years of age (mean of 7.6 years, 11 female), and 16 adults between 19.5 and 21.1 years of age (mean of 22.7 years, 12 female). Two additional children (5.2 and 10.6 years) were excluded based on their inability to detect AM even in the absence of a masker; this atypical result could reflect failure to understand the task or poor motivation.

The task was to detect sinusoidal AM on a 4300-Hz pure-tone carrier with AM rates of either 16, 64 or 256 Hz. The decision to concentrate on rates of 16 Hz and above represents a compromise between competing considerations: (1) the importance of low-rate envelope cues for speech perception (Drullman *et al.*, 1994b; Shannon *et al.*, 1995), (2) the need to avoid very low rates to allow effective estimates of modulation tuning and avoid effects related to a small number of AM periods (Sheft and Yost, 1990), (3) the contribution of high-rate envelope cues for speech perception (Stone *et al.*, 2008), and (4) the need to avoid introducing spectral cues associated with resolved sidebands (Stellmack *et al.*, 2005). The target tone was gated on for 500 ms, including 50-ms raised-cosine ramps. Thresholds were measured at all three AM rates without an off-frequency modulation masker. In addition, thresholds for 16-Hz AM were measured in the presence of a 1000-Hz masker tone that was sinusoidally modulated at 16 Hz with a modulation depth of 50%. That off-frequency masker was either continuous, or it was gated synchronously with the target. The target and the masker, when present, each had a level of 65 dB sound pressure level (SPL). The target modulation phase was randomly selected from a uniform distribution (0–2π) prior to each trial. The starting phase of the masker AM was randomly selected at the outset of a threshold estimation track. In addition to the tonal carriers, there was a continuous noise, lowpass filtered at 600 Hz (fourth-order biquad) and presented at 40-dB spectrum level in all conditions. The purpose of this noise was to prevent AM detection based on low-frequency distortions.

Most listeners took two 1-h sessions to complete this experiment, although some of the youngest children required a third session. Testing began for AM detection either with or without the off-frequency modulation masker. All conditions of this type were completed in quasi-random order before moving on to conditions of the other type. Due to the maximum target modulation depth of 100%, thresholds underestimated the level associated with 71% correct for some listeners in the more difficult listening conditions.[2]

B. Results and discussion

Analysis of the results focuses on two main questions. First, how does detection of AM on a pure-tone carrier depend on AM rate and listener age? Second, how does the introduction of an off-frequency masker affect AM detection?

1. Detection of AM as a function of AM rate

Figure 1 shows results in the three conditions without an off-frequency modulation masker for individual listeners, shown separately for target AM rates of 16, 64, and 256 Hz. Thresholds for child listeners are plotted as a function of age, and the distribution of thresholds for adults is shown at the right of each panel. Thresholds tended to improve as a function of child age. That trend was significant for the 256-Hz rate ($r = -0.42$, $p = 0.037$) and the 64-Hz rate ($r = -0.41$, $p = 0.040$), but not for the 16-Hz rate ($r = -0.17$, $p = 0.174$). Child data for these three conditions were analyzed using a linear mixed model with factors age and target AM rate. There was a significant effect of target AM rate ($F_{2,46} = 4.84$, $p = 0.012$). The effect of age ($F_{1,23} = 0.65$, $p = 0.429$) and the interaction between age and target AM rate ($F_{2,46} = 1.57$, $p = 0.218$) were not significant.

Although effects of child age for AM detection without an off-frequency modulation masker were weak and non-significant when evaluated in a single model, effects of development were significant in the analysis of group data. A linear mixed model comparing AM detection thresholds for children and adults, with factors age group and target AM rate, resulted in a main effect of target AM rate ($F_{2,78} = 5.97$, $p = 0.004$), a non-significant trend for a main effect of age group ($F_{1,39} = 3.39$, $p = 0.073$), and a significant interaction between target AM rate and age group

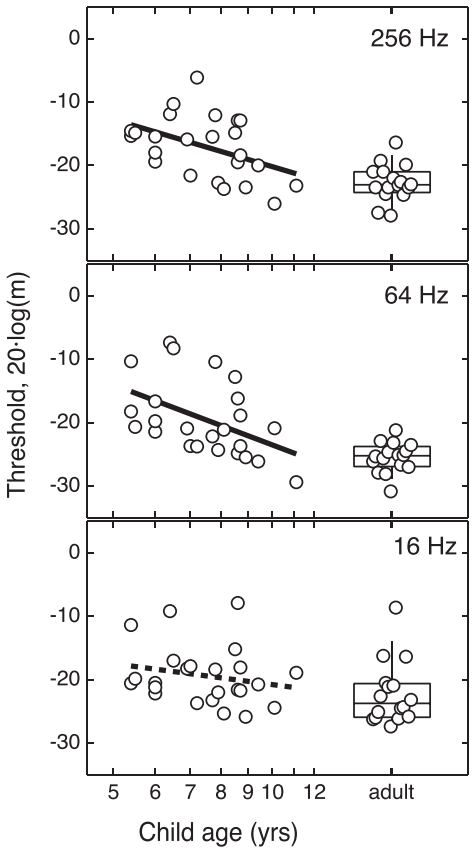

FIG. 1. Thresholds for detecting AM on a pure-tone carrier presented without an off-frequency modulation masker. Results for child listeners are plotted as a function of age. Solid lines indicate significant correlations between thresholds and child age, and the dotted line indicates a linear trend that failed to reach significance. The distribution of thresholds for adult listeners is shown at the far right of each panel. Horizontal lines indicate the median, boxes span the 25th–75th percentiles, and vertical lines span the 10th–90th percentiles.

($F_{2,78} = 3.27$, $p = 0.043$). This interaction remained significant when either the 64-Hz or 256-Hz data were removed from the model ($p \leq 0.043$) but not when 16-Hz data were removed from the model ($p = 0.901$). This indicates that the rate-by-group interaction in the full model was driven primarily by the fact that the child-adult difference was smaller for 16-Hz AM detection (mean = 2.7 dB) than for 64-Hz (mean = 5.8 dB) and 256-Hz AM detection (mean = 5.6 dB).

One possible explanation for children's relatively adult-like sensitivity to 16-Hz AM without an off-frequency modulation masker is that some children had prior experience with 16-Hz AM detection in the context of the masked AM conditions. This possibility is not supported by the data, however. Including condition order in the linear model evaluating performance by child age (masked conditions before vs after no-masker conditions) did not yield a significant effect ($F_{1,22} < 0.01$, $p = 0.964$). Another possibility is that variability in adult's 16-Hz AM thresholds reduced the child-adult difference.

The observation of relatively mature detection of 16-Hz AM is consistent with the results of Buss *et al.* (2013). In that study, AM detection thresholds were measured for 16-Hz sinusoidal AM carried on a 1000-Hz pure tone that was presented at either 35 or 75 dB SPL. The carrier tone was continuous, and AM was present in one of three listening intervals. Thresholds for 71% correct were approximately 10 dB better at 75 than 35 dB SPL, and that level effect was not different for children and adults. As in the present study, improvement in thresholds between 5.4 and 9.5 years of age was not significant, but there was a significant difference between thresholds for children and adults. Based on line fits to the published data, the child-adult differences at 5 years of age were 5.3 dB (35 dB SPL) and 4.0 dB (75 dB SPL). These results are consistent with the child-adult difference of 4.2 dB for 16-Hz AM detection with a 65-dB-SPL stimulus in the present experiment.

One motivation for measuring AM detection for a pure-tone carrier without an off-frequency modulation masker was to evaluate temporal resolution in the absence of modulation masking. Hall and Grose (1994) used a bandpass noise carrier to evaluate AM detection thresholds at 5, 20, 100, 150, and 200 Hz. Thresholds were higher for children than adults overall, but that difference did not depend on the AM rate. Across modulation rates, the child-adult difference was approximately 4.0–5.9 dB for 4- to 5-year-olds and 1.7–2.0 dB for 9- to 10-year-olds. The present study also found a child-adult difference in overall sensitivity to AM with a pure-tone carrier, but in contrast to the noise-carrier TMTF, there *was* evidence of a rate-by-age interaction with the pure-tone carrier. The child-adult difference for 16-Hz AM detection (4.2 dB at 5 years) was similar to that observed by Hall and Grose (1994) with a noise carrier, but the child-adult difference for 64- and 256-Hz AM detection (10.5 and 9.5 dB, respectively, at 5 years) was larger than that observed previously. These results are consistent with the conclusion that developmental effects for AM detection with a pure-tone carrier are more pronounced at 64 and 256 Hz than at 16 Hz. This possibility contrasts with the conclusions of a recent study by Cabrera *et al.* (2017). That study measured detection of AM at rates of 4, 8, and 32 Hz applied to a 1000-Hz pure-tone carrier, and found consistent child-adult differences irrespective of AM rate. One possibility is that development of temporal resolution is more evident at higher rates of AM (64 and 256 Hz) than lower rates (up to 32 Hz).

There is some indication in the literature that different cues underlie adults' low- and high-rate AM detection. For example, Wright and Dai (1998) tested adults' ability to detect AM at expected and unexpected rates, and they found that the prior expectation was more beneficial for low-rate than high-rate AM detection. This result was explained in terms of the different cues used by adults to detect AM at different rates: fluctuation in loudness for rates below 20 Hz and roughness or pitch for higher rates. They further argued that roughness is a more perceptually salient stimulus feature for adults than low-rate level fluctuation. The finding of a marked developmental effect for 64- and 256-Hz AM, but not 16-Hz AM, is consistent with more pronounced maturation in the ability to use the roughness cue than the level fluctuation cue for AM detection.

The distinction between loudness fluctuation and roughness cues is reminiscent of the distinction between template matching (Dau *et al.*, 1997a,b) and envelope power (Ewert and Dau, 2000), two decision statistics used in current

implementations of the AM filterbank model. In the *perception* model, the decision module computes the cross correlation between the internal representation of the current stimulus and an internal template stored in memory. This implementation of the model is successful at predicting low-rate AM detection thresholds (<5–10 Hz), provided that envelope phase is preserved at the output of modulation filters. In the *envelope power spectrum* model, the decision module computes the long-term envelope power at the output of modulation filters, discarding envelope phase information. This model is successful at predicting AM detection thresholds for rates above 5–10 Hz. The present results may therefore reflect distinct maturation of decision mechanisms operating on slow and fast envelope cues and, more speculatively, the development of temporal integration capacities. One caveat is that 16-Hz AM, the lowest rate tested in the present experiment, is higher than the range of rates thought to be better characterized by the template cue than the envelope power cue.

2. Detection of AM with off-frequency modulation maskers

Overall, thresholds for the continuous off-frequency modulation masker were similar to those measured without an off-frequency masker. This association is shown in Fig. 2. Analyzing data for children and adults separately, the correlation between 16-Hz AM thresholds with and without the continuous modulation masker was $r = 0.81$ ($p < 0.001$) in both cases. Thresholds were not significantly higher with than without the off-frequency masker for either children (mean = 0.69, $t_{24} = 1.15$, $p = 0.261$) or adults (mean = 0.15, $t_{15} = 0.21$, $p = 0.840$). This indicates that there was essentially no modulation masking in the continuous off-frequency masker condition.

Figure 3 shows 16-Hz AM detection thresholds measured with the continuous and gated off-frequency modulation maskers (top and middle panels, respectively), and the difference between thresholds in these two conditions (bottom panel). In the child data, thresholds improved with

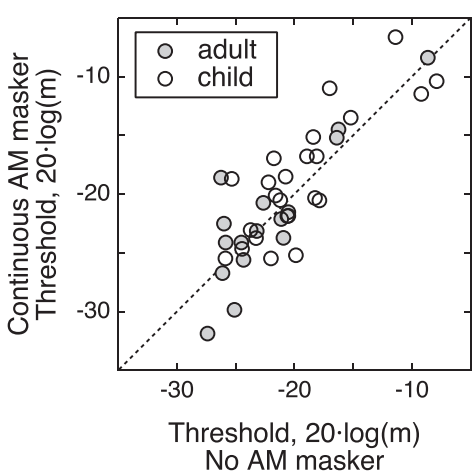

FIG. 2. Thresholds for detecting 16-Hz AM with a continuous off-frequency modulation masker, plotted as a function of the associated 16-Hz AM threshold measured with no modulation masker. Filled symbols indicate data for adults, and open symbols indicate data for children. The diagonal line indicates a perfect match.

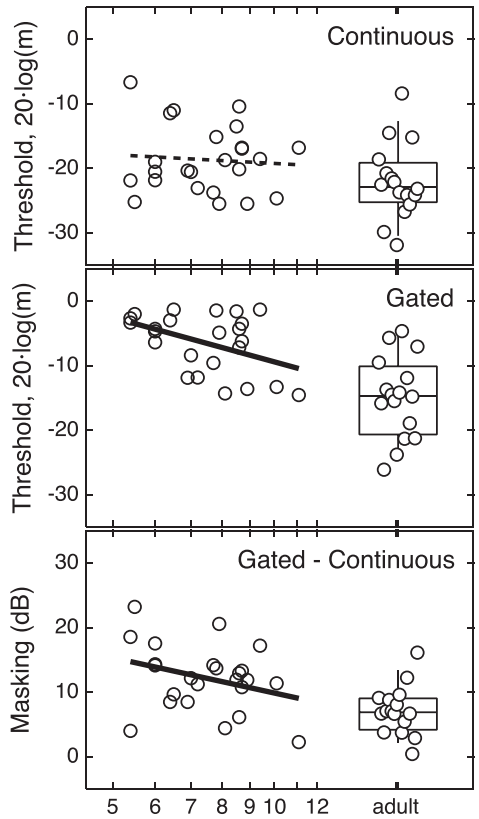

FIG. 3. Results obtained with the off-frequency modulation masker. Thresholds are plotted for individual listeners tested with the continuous masker (top) and the gated masker (middle), and differences between thresholds in these two conditions are shown (bottom). Plotting conventions follow those of Fig. 1.

increasing age for the gated masker ($r = -0.45$, $p = 0.025$) but not for the continuous masker ($r = -0.04$, $p = 0.849$). The child data for these two conditions were evaluated in a linear mixed model with factors age and stimulus condition. There was a significant effect of condition ($F_{1,23} = 13.53$, $p = 0.001$), no main effect of age ($F_{1,23} = 0.06$, $p = 0.816$), and a non-significant trend for an interaction between condition and age ($F_{1,23} = 3.16$, $p = 0.089$). An analysis of group data for children and adults indicated a main effect of condition ($F_{1,39} = 178.06$, $p < 0.001$), a non-significant trend for a main effect of age group ($F_{1,39} = 3.62$, $p = 0.064$), and a significant interaction between condition and age group ($F_{1,39} = 12.48$, $p = 0.001$). This interaction reflects the fact that the child-adult difference was smaller for the continuous masker (mean = 3.3 dB; $t_{39} = 1.87$, $p = 0.068$) than for the gated masker (mean = 8.5 dB; $t_{39} = 5.05$, $p < 0.001$). The interaction between condition and age is illustrated in the bottom panel of Fig. 3, which shows a larger difference between conditions for children than adults. Based on line fits, the magnitude of this child-adult difference was 8.7 dB for 5-year-olds, although this value is likely to be an underestimate due to ceiling effects in the gated off-frequency masker condition.

For both age groups, effects of the off-frequency masker depended on whether it was continuous or gated on and off with the target. When the masker was continuous, neither children nor adults experienced masking. A difference between

results obtained with continuous and gated off-frequency maskers is often interpreted as reflecting auditory stream segregation. Whereas synchronous gating tends to promote fusion of the target and masker into a single auditory object (Shamma *et al*., 2011), target onset in the context of a continuous masker facilitates segregation and auditory attention to the target frequency (Oxenham and Dau, 2001). The mechanism of this benefit could be related to the exogenous attention capture associated with onsets or to adaptation to the masker stimulus (Jonides and Yantis, 1988; Dalton and Lavie, 2007). Most previous data from adults indicate some residual masking when the target is gated and the masker is continuous (Hall and Grose, 1991; Moore and Shailer, 1992; Oxenham and Dau, 2001), although there is precedent in the literature for elimination of off-frequency modulation masking using a sequential streaming manipulation (Oxenham and Dau, 2001). The finding that neither children nor adults exhibited masking in the continuous masker condition suggests that children are able to use gating asynchrony as a segregation cue to a similar degree as adults, an ability that has been observed in other developmental datasets (Leibold and Neff, 2007).

In contrast to the continuous masker, the gated off-frequency masker did elevate thresholds relative to the unmasked case for listeners of all ages. Threshold elevation associated with introducing the gated masker was more pronounced for younger children than for older children and adults. For adults, the average difference between thresholds with the gated and continuous off-frequency maskers was 7.1 dB. This is similar to the effect of 6–11 dB previously reported for adults for low-rate AM detection in the presence of a synchronously gated off-frequency masker (Yost *et al*., 1989; Hall and Grose, 1991). Modulation masking for 5-year-olds was estimated to be at least twice that observed for adults. The finding that young children are more susceptible to off-frequency modulation masking than adults is consistent with results of other paradigms showing that the ability to listen in a frequency-selective manner develops with age (Werner and Bargones, 1991; Leibold and Neff, 2011; Leibold and Buss, 2016; Youngdahl *et al*., 2018). Greater modulation masking in the gated off-frequency masker condition for children than adults likely reflects children's reduced ability to segregate and selectively attend to the target frequency under these conditions.

Although young children are more susceptible to off-frequency modulation masking than older children and adults, it is unlikely that this result explains the finding of an age-by-rate interaction for AM detection with a tone carrier (present dataset) but not a noise carrier (Hall and Grose, 1994). In contrast to the case of a target tone with an off-frequency masker, focusing auditory attention on a restricted carrier frequency is not an effective strategy for reducing masking when both the target and inherent masker modulation occur in the same carrier frequency regions. Thus, on-frequency modulation masking is more likely than off-frequency modulation masking to play a role in the noise-carrier TMTF result. On-frequency masking with a noise carrier is the topic of the second experiment.

IV. EXPERIMENT 2: AM DETECTION FOR A NOISE CARRIER WITH AND WITHOUT AN ON-FREQUENCY MODULATION MASKER

For adults, differences between TMTFs measured with noise carriers and pure-tone carriers have been attributed to modulation masking, which is associated with inherent fluctuation of the noise (Dau *et al*., 1997a,b; Dau *et al*., 1999). It is unclear what role modulation masking plays in the noise-carrier TMTF measured for school-age children, however, in part because modulation masking in this age group is poorly understood. The second experiment therefore evaluated on-frequency modulation masking and tuning to modulation masking in children and adults. Target modulation rates of 16 Hz and 64 Hz were examined, based on the observation in experiment 1 of differential developmental effects for these two rates with a pure-tone carrier. This experiment was modeled after Ewert and Dau (2000). In that study, adults detected sinusoidal AM carried on a band of noise. Masker AM was imposed via multiplication with a raised narrowband noise sample, and the overlap between the spectral content of target and masker AM was manipulated parametrically.

For adults, modulation masking is sensitive to the spectral similarity between target and masker AM with maximal masking for spectrally overlapping target and masker AM, and reduced masking when the modulation spectra differ (Ewert and Dau, 2000). This result is described as tuning to AM rate. One general prediction is that children should perform more poorly than adults overall, consistent with previous evidence of reduced efficiency in children. A second prediction is that children have broader tuning to modulation rate than adults. We know that young children are less adept than adults at selectively attending to the auditory filter associated with the target carrier frequency (Leibold and Buss, 2016). It is possible that selectively monitoring the modulation filter associated with the target AM relies on the same underlying abilities.

A. Methods

Different groups of listeners were recruited for the 16- and 64-Hz AM detection tasks. The first group comprised 20 children between 5.2 and 11.5 years of age (mean of 8.0 years, 8 female) and 10 adults between 22.1 and 42.5 years of age (mean of 27.8 years, 6 female). The second group comprised 23 children between 4.7 and 10.4 years of age (mean of 7.5 years, 17 female) and 8 adults between 20.0 and 44.4 years of age (mean of 28.4 years, 6 female).

The task was to detect sinusoidal AM imposed on a Gaussian noise carrier that was bandpass filtered using a first-order biquad filter. At the 10-dB-down points, this band spanned 1000–5000 Hz for 16-Hz AM detection and 2500–5500 Hz for 64-Hz AM detection. A higher-frequency carrier was used for the higher target rate to guard against spectral cues associated with AM. The stimulus was played continuously at 65 dB SPL. Target AM, when present, was 600 ms in duration, including 200-ms raised-cosine ramps. Thresholds were measured with and without continuous on-frequency masker AM. The masker AM, when present, was generated via multiplication with a raised band of low-fluctuation noise. For

the 16-Hz AM target, this band was 5.6 Hz wide, and it was centered on 6.3, 16, or 40.3 Hz. For the 64-Hz AM target, this band was 22.3 Hz wide, and it was centered on 25.4, 64, or 161.3 Hz. The resulting patterns of masker AM had nominal rates 1.33 oct below the target AM rate, at the target AM rate, or 1.33 oct above the target AM rate. Low-fluctuation noise bands were generated using an iterative procedure, whereby the time waveform was divided by its Hilbert envelope and restricted to its original bandwidth in the frequency domain, a process that was repeated eight times. A new 21-s noise sample was generated prior to each threshold estimation track for which masker modulation was present.

The method for estimating the AM detection threshold was modified from that used by Ewert and Dau (2000). In that study, masker AM depth was fixed, and target AM depth was adaptively varied. In the present study, the overall AM depth was fixed, and the signal-to-noise ratio (SNR) of the target AM depth and masker AM depth was adaptively varied. The rationale for this change was that the poorest-performing children might struggle to detect 100% target AM in the context of masker AM. To avoid ceiling effects, the combined target and masker AM was fixed at −7 dB relative to 100% AM depth.[3] Setting the combined AM depth below 0 dB guards against overmodulation. Changes in masker AM depth were implemented between trials and smoothed over 500 ms. The target was 16- or 64-Hz sinusoidal AM, imposed via multiplication with the noise carrier or the product of the noise carrier and masker AM envelope. The target and masker envelope functions were each scaled by $1 + m^2/2$ prior to multiplication with the carrier in order to prevent changes in presentation level with changes in AM depth.

B. Results

Analyses of results focused on two main questions. First, how does AM detection with a noise carrier compare with the pure-tone AM detection observed in experiment 1? Second, does modulation masking differentially affect children and adults?

1. Detection of AM without added masker AM

Thresholds without an AM masker are plotted for individual listeners in Fig. 4 with 64-Hz AM thresholds shown in the top panel and 16-Hz AM thresholds in the bottom panel. Data for child listeners are plotted as a function of age at the left, and the distribution of thresholds for adult listeners is plotted at the right in each panel. There was a correlation between threshold and child age for the 64-Hz AM ($r = -0.55$, $p = 0.007$) but not for the 16-Hz AM ($r = 0.05$, $p = 0.842$). A linear mixed model, including child age and target AM rate, indicated an effect of target rate ($F_{1,39} = 59.80$, $p < 0.001$), no effect of age ($F_{1,39} = 0.05$, $p = 0.826$), and a non-significant trend for an interaction between age and rate ($F_{1,39} = 3.49$, $p = 0.069$). There was a significant difference between thresholds for children and adults for both 64-Hz AM ($t_{29} = 3.88$, $p < 0.001$) and 16-Hz AM ($t_{28} = 3.39$, $p = 0.002$).

For adults, 16-Hz AM detection thresholds with the noise carrier were not significantly different from the mean value observed with a pure-tone carrier in experiment 1

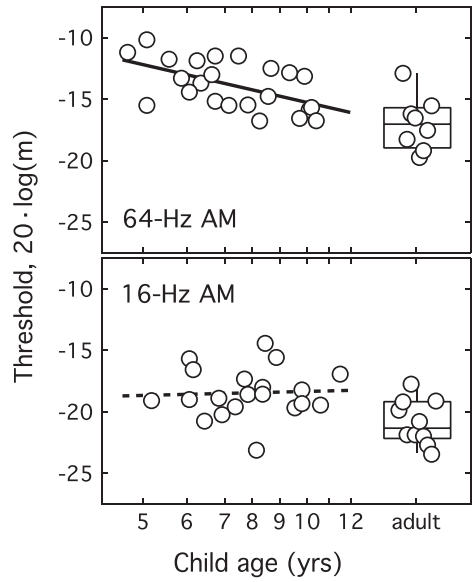

FIG. 4. Thresholds for detecting 64-Hz AM (top) or 16-Hz AM (bottom), carried on a bandpass noise without additional AM masking applied. Plotting conventions follow those of Fig. 1.

(−21.0 vs −22.2 dB; $t_{20.6} = 0.90$. $p = 0.381$), but thresholds were significantly higher for the noise than the pure-tone carrier at the 64-Hz rate (−17.1 vs −25.5; $t_{14.9} = 8.61$, $p < 0.001$). This result is consistent with significant modulation masking for 64-Hz AM detection even in the absence of additional masker AM.

One question of interest is whether the modulation masking seen for adults at 64 Hz was also seen for children. Comparing line fits to child data, the difference between thresholds with the pure-tone and noise carriers rose from 2.9 dB at 5 years to 8.8 dB at 11 years of age. However, the interaction between age and carrier type did not reach significance in a linear model of 64-Hz AM detection in children ($F_{1,44} = 1.86$, $p = 0.180$). If this data pattern is reliable, then it could reflect reduced effects of modulation masking due to immature temporal resolution or higher levels of internal noise. By this account, young children's reduced sensitivity to AM in the region of 64 Hz would make them less susceptible to modulation masking. Regardless, results of the present experiment are inconsistent with the hypothesis that younger children are *more* susceptible to on-frequency modulation masking than older children and adults.

2. Detection of AM with on-frequency masker AM

Figure 5 shows results for individual listeners in conditions incorporating masker AM. In all six conditions, there was a trend for thresholds to improve as a function of child age with correlations ranging between $r = -0.55$ ($p = 0.011$; 64-Hz target and 25.4-Hz masker) and $r = -0.16$ ($p = 0.507$; 16-Hz target and 40.3-Hz masker). A linear mixed model was fitted to child data, including factors child age, target AM rate, and relative masker AM (below, at, or above the target AM rate). A model without interactions indicated significant effects of all three factors, including age ($p = 0.018$). A model that included two- and three-way interactions indicated significant effects of target AM rate ($F_{1,39} = 8.21$,

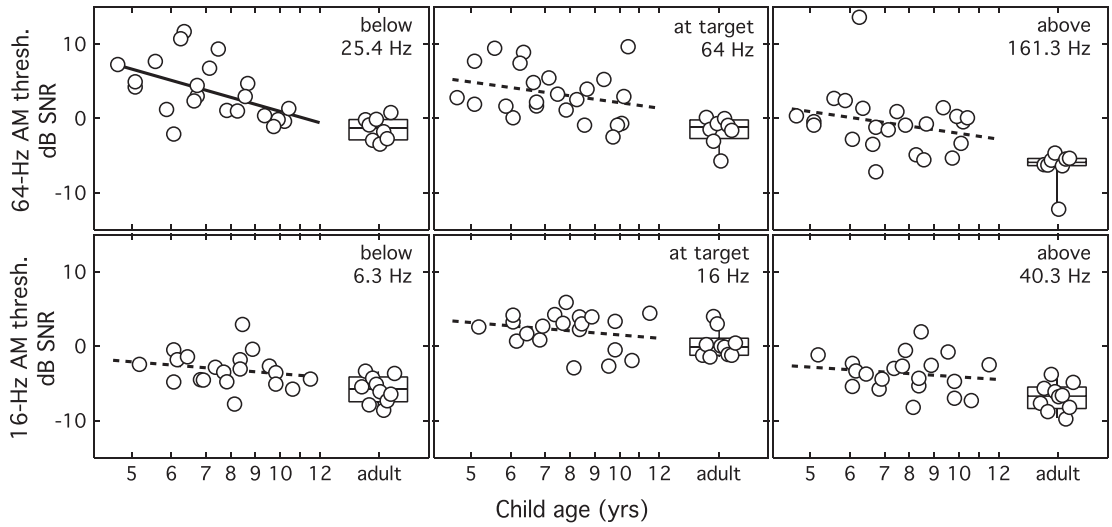

FIG. 5. Thresholds for detecting 64-Hz AM (top) or 16-Hz AM (bottom) carried by a bandpass noise in dB SNR relative to masker AM with a fixed overall AM depth of −7 dB. Data are shown separately for masker modulation rates below the target rate (left), at the target rate (center), and above the target rate (right). Values in the upper right of each panel indicate the nominal masker AM rate. Plotting conventions within each panel follow those of Fig. 1.

$p = 0.007$) and relative masker AM rate ($F_{2,78} = 38.96$, $p < 0.001$) and a significant interaction between target AM rate and relative masker AM rate ($F_{2,78} = 15.10$, $p < 0.001$); however, there was no significant effect of age or interaction with age ($p \geq 0.586$). Repeating this analysis for thresholds from both adults and children, with age as a group variable, there was a significant effect of age group ($F_{1,57} = 8.89$, $p = 0.004$), a significant effect of relative masker AM rate ($F_{2,114} = 37.72$, $p < 0.001$), and a significant interaction between target AM rate and relative masker AM rate ($F_{2,114} = 11.81$, $p < 0.001$). Of particular importance here, none of the interactions with age group were significant ($p \geq 0.112$). The main effect of age group is consistent with reduced efficiency in children compared to adults. The effect of relative masker AM rate and the lack of an interaction with age indicates consistent tuning to masker AM rate.

Thresholds of Ewert and Dau (2000) cannot be directly compared to those collected in experiment 2 because the published study adaptively varied target AM depth, and the present study varied the target-to-masker ratio. However, the general pattern of results can be compared. In Ewert and Dau (2000), thresholds improved as the masker AM rate deviated from the target AM rate, but threshold improvements for equal deviations on a log scale were larger for 16-Hz than 64-Hz target AM. Comparing thresholds for deviations of ±1.33 oct, the deviations used in the present study, thresholds improved by means of 9.2 dB for the 16-Hz AM target and 4.4 dB for the 64-Hz AM target. For both target AM rates, the benefit of moving the masker rate away from the target rate was smaller for masker AM below the target AM rate than above the target AM rate. For the 16-Hz target AM in the present study, thresholds were lower for masker AM rates below and above the target rate, with differences of 5.2 and 6.2 dB, respectively. In contrast, there was considerable asymmetry in the effect observed for 64-Hz target AM. Compared to 64-Hz masker AM, thresholds improved by 6.7 dB when masker AM was above the target AM rate, but no benefit was observed when the masker AM was below the target rate. While it is unclear how to explain a failure to benefit from offsetting the masker AM by −1.33 oct, asymmetry of a comparable magnitude is evident for a target-masker AM separation of 0.66 oct in the data of Ewert and Dau (2000).

V. GENERAL DISCUSSION

The two experiments reported here evaluated different aspects of AM detection and modulation masking for normal-hearing school-age children and adults. The first experiment measured detection thresholds for three different rates of sinusoidal AM carried on a 4300-Hz pure tone. Children performed more poorly than adults overall, and the child-adult difference was larger for 64- and 256-Hz AM than for 16-Hz AM. This finding suggests children may have poorer temporal resolution for high-rate AM than adults. Failure to observe a rate-by-age interaction in the noise-carrier TMTF paradigm (Hall and Grose, 1994) could be due to modulation masking. The first experiment also showed that an off-frequency masker, which was gated on and off with the target tone, elevated AM thresholds more for younger children than for older children and adults; none of the groups experienced masking when the off-frequency masker was played continuously. A child-adult difference with the gated masker but not with the continuous masker implicates a failure of auditory stream segregation, a failure of selective attention in children's susceptibility to off-frequency modulation masking, or a combination of these factors. The second experiment assessed the effects of on-frequency modulation masking. That dataset provided evidence of a modest child-adult difference in all conditions, including those with and without masker modulation. However, there was no evidence of a child-adult difference in rate tuning or greater susceptibility to on-frequency modulation masking for children. In fact, there was a non-significant trend in the opposite direction—for less modulation masking in children than adults.

Previous data for adults show that whereas noise-carrier TMTFs begin to roll off at 50 Hz, tone-carrier TMTFs begin to roll off at 100–130 Hz. This difference in shape implies that modulation masking has a greater effect for AM at rates of 50 Hz and above. The previous finding of no age effect in the shape of the noise TMTF could therefore be due to modulation masking obscuring group differences in temporal resolution at 50-Hz AM and above. This possibility is undermined somewhat by the observation in experiment 2 of a non-significant trend for an age effect for 64-Hz AM detection but not 16-Hz AM detection with a noise carrier.

Although on- and off-frequency modulation masking were evaluated using different paradigms in the present report, it is unlikely that these differences are responsible for the different results. The masker AM was sinusoidal in experiment 1, but it was a bandpass noise in experiment 2. While previous studies of off-frequency modulation masking have typically used sinusoidal AM maskers (Hall and Grose, 1991; Moore and Shailer, 1992), similar results are obtained when masker envelopes are raised noise samples (Mendoza *et al*., 1995; Oxenham and Dau, 2001). Another difference between experiments is the gating applied to the target carrier: the carrier was gated in experiment 1 but continuous in experiment 2. Oxenham and Dau (2001) showed that off-frequency modulation masking could be demonstrated with a continuous carrier, and there is no reason to expect a gating effect for on-frequency modulation masking.

It has been proposed that speech-in-noise recognition in adults may be affected by both on-frequency (Stone *et al*., 2011; Stone *et al*., 2012; Stone and Moore, 2014) and off-frequency modulation masking (Apoux and Bacon, 2008). This has led to speculation that susceptibility to modulation masking in children could play a role in the child-adult difference observed for speech in noise (Buss *et al*., 2018). In the present study, there was evidence of a child-adult difference for 16-Hz AM detection with an off-frequency masker (experiment 1) but not with an on-frequency masker (experiment 2). While there could be some effect of off-frequency masking for speech in noise, this effect is likely to be small for two reasons. First, the target AM associated with the speech signal occurs in frequency channels that also represent inherent noise fluctuation (masker AM), so selective attention to a particular frequency region would not eliminate modulation masking. Second, it has been argued that both children and adults are able to segregate speech from a noise masker (Freyman *et al*., 1999; Leibold *et al*., 2016). If segregation of the target and masker substantially reduces or eliminates off-frequency modulation masking for both children and adults, then we would not expect off-frequency modulation masking to play an important role in speech-in-noise recognition. In contrast to speech recognition in a noise masker, auditory stream segregation is a challenge for both children and adults for speech in a speech masker (Leibold *et al*., 2016; Corbin *et al*., 2017). One question for future research is whether off-frequency modulation masking plays a role in the marked child-adult difference observed for speech-in-speech recognition.

Results of a recent study designed to evaluate the effect of modulation masking for speech in noise recognition by children and adults failed to find support for the idea that modulation masking is responsible for the child-adult difference for speech-in-noise recognition. Buss *et al*. (2018) measured masked sentence recognition for narrow bands of speech masked by either spectrally overlapping bands of noise or pure tones centered on each speech band. As previously observed by Stone *et al*. (2012), adults' thresholds were lower for the multi-tone masker than for the multi-noise-band masker. This difference has been interpreted as reflecting the modulation masking associated with inherent AM of the noise bands. If children experience more modulation masking than adults, then the child-adult difference should be larger for the multi-noise-band masker than the multi-tone masker. However, the opposite effect was observed—there were *smaller* child-adult differences for the multi-noise-band than for the multi-tone masker.

Buss *et al*. (2018) argued that the large child-adult difference for the multi-tone masker reflects children's reduced ability to selectively attend to the frequency regions associated with the best SNR. This effect would be seen in the multi-tone but not the multi-noise-band masker conditions due to the greater differences in SNR across frequency in the former than the latter condition. This interpretation—in terms of the ability to selectively attend to the most informative frequency regions—is consistent with the data on children and adults in the off-frequency modulation masking paradigm (experiment 1). The most parsimonious explanation for both results is that segregation and selective attention to the target, not on-frequency modulation masking, limit children's performance.

While modulation masking may not play an important role in the child-adult difference for speech-in-noise recognition, results obtained for AM detection without an off-frequency masker indicate that temporal resolution could be an important factor. Low- and high-rate AM provide different types of speech information (Rosen, 1992): envelope cues between 2 and 50 Hz reflect movement of the articulators, whereas periodicity cues in the range of 50–500 Hz convey information about voicing and voice pitch. Speech can be recognized in quiet based on low-rate AM cues (Drullman *et al*., 1994b; Shannon *et al*., 1995) and associated temporal fine-structure cues (Shamma and Lorenzi, 2013). However, higher-rate cues become more important when maskers are present, particularly complex maskers like speech (Qin and Oxenham, 2003, 2005; Stone *et al*., 2008). This suggests that the effects of temporal resolution on speech perception in children could depend on the characteristics of the masker. Children may be at a disadvantage under challenging listening conditions due to the increased importance of high-rate cues, the very cues that they are less adept than adults at using. For example, a recent study by Flaherty *et al*. (2018) reported that whereas adults benefit from an $F0$ difference between the target and masker speech, young children derive little or no benefit from $F0$ mismatches. One possible factor contributing to this result is that children's limited temporal resolution for high-rate envelope cues interferes with their ability to utilize segregation cues based on differences in voice $F0$.

ACKNOWLEDGMENTS

This work was supported by National Institutes of Health (NIH) National Institute on Deafness and Other Communication Disorders (NIDCD) Grant No. R01DC000397.

[1]Initial steps for adjusting the AM depth were 4 dB for the pure-tone carrier without an off-frequency masker (experiment 1) and 8 dB for all other conditions, including the noise carrier (experiment 2) and the pure-tone carrier with an off-frequency masker (experiment 1). The rationale for using larger initial steps in the latter conditions was to ensure that the listener had the opportunity to hear supra-threshold trials under conditions of modulation masking.

[2]Logit functions were fitted to adaptive track data in an attempt to reduce the bias associated with the upper limit of target AM depth. Thresholds based on function fits were correlated with those computed as the mean of reversals ($r = 0.93$). However, the quality of the fits to some individuals' data was quite poor. Fits with $r^2 < 25\%$ were most common for the gated off-frequency masker condition, occurring for 2/16 adults and 10/25 children. One factor contributing to poor fits could be poor performance even at the ceiling AM depth of $m = 1$. Adaptive tracks reached $m = 1$ more than three times for 1–2 children in most conditions; the one exception was the gated off-frequency masker condition for which the ceiling was reached more than three times by 14/25 children. This result suggests that children's thresholds in the gated masker condition likely overestimated sensitivity to the target AM. In no case did an adult's adaptive track hit the ceiling.

[3]Fixing the combined target and masker AM depth at −7 dB was accomplished as follows: $T_{\text{scaled}} = T - 10\log_{10}(10^{M/10} + 10^{T/10}) - 7$ and $M_{\text{scaled}} = M - 10\log_{10}(10^{M/10} + 10^{T/10}) - 7$. Variables T and M are the nominal modulation depths of target AM (dB SNR) and masker AM (0-dB reference). The variables T_{scaled} and M_{scaled} are the modulation depths used in stimulus generation.

Apoux, F., and Bacon, S. P. (**2008**). "Selectivity of modulation interference for consonant identification in normal-hearing listeners," J. Acoust. Soc. Am. **123**, 1665–1672.

Bacon, S. P., and Grantham, D. W. (**1989**). "Modulation masking: Effects of modulation frequency, depth, and phase," J. Acoust. Soc. Am. **85**, 2575–2580.

Buss, E., Hall, J. W., 3rd, and Grose, J. H. (**2013**). "Factors affecting the processing of intensity in school-aged children," J. Speech Lang. Hear. Res. **56**, 71–80.

Buss, E., Leibold, L. J., and Lorenzi, C. (**2018**). "Speech recognition for school-age children and adults tested in multi-tone vs multi-noise-band maskers," J. Acoust. Soc. Am. **143**, 1458–1466.

Cabrera, L., Rosen, S., Buss, E., and Lorenzi, C. (**2017**). "Modulation masking and temporal integration for amplitude modulation in childhood," in *The 40th Midwinter Research Meeting of the Association for Research in Otolaryngology*, Baltimore, MD.

Corbin, N., Bonino, A. Y., Buss, E., and Leibold, L. J. (**2016**). "Development of open-set word recognition in children: Speech-shaped noise and two-talker speech maskers," Ear. Hear. **37**, 55–63.

Corbin, N. E., Buss, E., and Leibold, L. J. (**2017**). "Spatial release from masking in children: Effects of simulated unilateral hearing loss," Ear. Hear. **38**, 223–235.

Dalton, P., and Lavie, N. (**2007**). "Overriding auditory attentional capture," Percept. Psychophys. **69**, 162–171.

Dau, T., Kollmeier, B., and Kohlrausch, A. (**1997a**). "Modeling auditory processing of amplitude modulation. I. Detection and masking with narrow-band carriers," J. Acoust. Soc. Am. **102**, 2892–2905.

Dau, T., Kollmeier, B., and Kohlrausch, A. (**1997b**). "Modeling auditory processing of amplitude modulation. II. Spectral and temporal integration," J. Acoust. Soc. Am. **102**, 2906–2919.

Dau, T., Verhey, J., and Kohlrausch, A. (**1999**). "Intrinsic envelope fluctuations and modulation-detection thresholds for narrow-band noise carriers," J. Acoust. Soc. Am. **106**, 2752–2760.

Drullman, R., Festen, J. M., and Plomp, R. (**1994a**). "Effect of reducing slow temporal modulations on speech reception," J. Acoust. Soc. Am. **95**, 2670–2680.

Drullman, R., Festen, J. M., and Plomp, R. (**1994b**). "Effect of temporal envelope smearing on speech reception," J. Acoust. Soc. Am. **95**, 1053–1064.

Ewert, S. D., and Dau, T. (**2000**). "Characterizing frequency selectivity for envelope fluctuations," J. Acoust. Soc. Am. **108**, 1181–1196.

Flaherty, M. M., Buss, E., and Leibold, L. J. (**2018**). "Developmental effects in children's ability to benefit from *F*0 differences between target and masker speech," Ear. Hear. (published online).

Freyman, R. L., Helfer, K. S., McCall, D. D., and Clifton, R. K. (**1999**). "The role of perceived spatial separation in the unmasking of speech," J. Acoust. Soc. Am. **106**, 3578–3588.

Grimault, N., Bacon, S. P., and Micheyl, C. (**2002**). "Auditory stream segregation on the basis of amplitude-modulation rate," J. Acoust. Soc. Am. **111**, 1340–1348.

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.

Hall, J. W., III, and Grose, J. H. (**1991**). "Some effects of auditory grouping factors on modulation detection interference (MDI)," J. Acoust. Soc. Am. **90**, 3028–3035.

Houtgast, T. (**1989**). "Frequency selectivity in amplitude-modulation detection," J. Acoust. Soc. Am. **85**, 1676–1680.

Jonides, J., and Yantis, S. (**1988**). "Uniqueness of abrupt visual onset in capturing attention," Percept. Psychophys. **43**, 346–354.

Kohlrausch, A., Fassel, R., and Dau, T. (**2000**). "The influence of carrier level and frequency on modulation and beat-detection thresholds for sinusoidal carriers," J. Acoust. Soc. Am. **108**, 723–734.

Leibold, L. J. (**2012**). "Development of auditory scene analysis and auditory attention," in *Human Auditory Development* (Springer, New York), pp. 137–161.

Leibold, L. J., and Buss, E. (**2016**). "Factors responsible for remote-frequency masking in children and adults," J. Acoust. Soc. Am. **140**, 4367–4377.

Leibold, L. J., and Neff, D. L. (**2007**). "Effects of masker-spectral variability and masker fringes in children and adults," J. Acoust. Soc. Am. **121**, 3666–3676.

Leibold, L. J., and Neff, D. L. (**2011**). "Masking by a remote-frequency noise band in children and adults," Ear Hear. **32**, 663–666.

Leibold, L. J., Yarnell Bonino, A., and Buss, E. (**2016**). "Masked speech perception thresholds in infants, children, and adults," Ear Hear. **37**, 345–353.

McCreery, R. W., Spratford, M., Kirby, B., and Brennan, M. (**2017**). "Individual differences in language and working memory affect children's speech recognition in noise," Int. J. Audiol. **56**, 306–315.

Mendoza, L., Hall, J. W., and Grose, J. H. (**1995**). "Modulation detection interference using random and sinusoidal amplitude modulation," J. Acoust. Soc. Am. **97**, 2487–2492.

Moore, B. C. J., and Shailer, M. J. (**1992**). "Modulation discrimination interference and auditory grouping," Philos. Trans. R. Soc. London B Biol. Sci. **336**, 339–346.

Nie, Y. J., and Nelson, P. B. (**2015**). "Auditory stream segregation using amplitude modulated bandpass noise," Front. Psychol. **6**(1151), 1–11.

Oxenham, A. J., and Dau, T. (**2001**). "Modulation detection interference: Effects of concurrent and sequential streaming," J. Acoust. Soc. Am. **110**, 402–408.

Park, M. H., Won, J. H., Horn, D. L., and Rubinstein, J. T. (**2015**). "Acoustic temporal modulation detection in normal-hearing and cochlear implanted listeners: Effects of hearing mechanism and development," J. Assoc. Res. Otolaryngol. **16**, 389–399.

Qin, M. K., and Oxenham, A. J. (**2003**). "Effects of simulated cochlear-implant processing on speech reception in fluctuating maskers," J. Acoust. Soc. Am. **114**, 446–454.

Qin, M. K., and Oxenham, A. J. (**2005**). "Effects of envelope-vocoder processing on *F*0 discrimination and concurrent-vowel identification," Ear Hear. **26**, 451–460.

Rashid, M. S., Leensen, M. C., and Dreschler, W. A. (**2016**). "Application of the online hearing screening test 'Earcheck': Speech intelligibility in noise in teenagers and young adults," Noise Health **18**, 312–318.

Rosen, S. (**1992**). "Temporal information in speech: Acoustic, auditory and linguistic aspects," Philos. Trans. R. Soc. London B Biol. Sci. **336**, 367–373.

Sek, A., Baer, T., Crinnion, W., Springgay, A., and Moore, B. C. (**2015**). "Modulation masking within and across carriers for subjects with normal and impaired hearing," J. Acoust. Soc. Am. **138**, 1143–1153.

Shamma, S. A., Elhilali, M., and Micheyl, C. (**2011**). "Temporal coherence and attention in auditory scene analysis," Trends Neurosci. **34**, 114–123.

Shamma, S., and Lorenzi, C. (**2013**). "On the balance of envelope and temporal fine structure in the encoding of speech in the early auditory system," J. Acoust. Soc. Am. **133**, 2818–2833.

Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., and Ekelid, M. (**1995**). "Speech recognition with primarily temporal cues," Science **270**, 303–304.

Sheft, S., and Yost, W. A. (**1990**). "Temporal integration in amplitude modulation detection," J. Acoust. Soc. Am. **88**, 796–805.

Stellmack, M. A., Viemeister, N. F., and Byrne, A. J. (**2005**). "Monaural and interaural temporal modulation transfer functions measured with 5-kHz carriers," J. Acoust. Soc. Am. **118**, 2507–2518.

Stone, M. A., Fullgrabe, C., Mackinnon, R. C., and Moore, B. C. J. (**2011**). "The importance for speech intelligibility of random fluctuations in 'steady' background noise," J. Acoust. Soc. Am. **130**, 2874–2881.

Stone, M. A., Fullgrabe, C., and Moore, B. C. J. (**2008**). "Benefit of high-rate envelope cues in vocoder processing: Effect of number of channels and spectral region," J. Acoust. Soc. Am. **124**, 2272–2282.

Stone, M. A., Fullgrabe, C., and Moore, B. C. J. (**2012**). "Notionally steady background noise acts primarily as a modulation masker of speech," J. Acoust. Soc. Am. **132**, 317–326.

Stone, M. A., and Moore, B. C. J. (**2014**). "On the near non-existence of 'pure' energetic masking release for speech," J. Acoust. Soc. Am. **135**, 1967–1977.

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.

Verhey, J., and Dau, T. (**2000**). "Modulation detection interference with narrow-band noise interferers," J. Acoust. Soc. Am. **107**, 2915.

Viemeister, N. F. (**1979**). "Temporal modulation transfer functions based upon modulation thresholds," J. Acoust. Soc. Am. **66**, 1364–1380.

Werner, L. A., and Bargones, J. Y. (**1991**). "Sources of auditory masking in infants: Distraction effects," Percept. Psychophys. **50**, 405–412.

Wright, B. A., and Dai, H. (**1998**). "Detection of sinusoidal amplitude modulation at unexpected rates," J. Acoust. Soc. Am. **104**, 2991–2996.

Yost, W. A., and Sheft, S. (**1989**). "Across-critical-band processing of amplitude-modulated tones," J. Acoust. Soc. Am. **85**, 848–857.

Yost, W. A., Sheft, S., and Opie, J. (**1989**). "Modulation interference in detection and discrimination of amplitude modulation," J. Acoust. Soc. Am. **86**, 2138–2147.

Youngdahl, C. L., Healy, E. W., Yoho, S. E., Apoux, F., and Holt, R. F. (**2018**). "The effect of remote masking on the reception of speech by young school-age children," J. Speech Lang. Hear. Res. **61**, 420–427.