# Journal Pre-proof

Temporal integration for amplitude modulation in childhood:
Interaction between internal noise and memory

Laurianne Cabrera ,  Irene Lorenzini ,  Stuart Rosen ,  Léo Varnet ,
Christian Lorenzi

Please cite this article as:  Laurianne Cabrera ,  Irene Lorenzini ,  Stuart Rosen ,  Léo Varnet ,  Christian Lorenzi , Temporal integration for amplitude modulation in childhood: Interaction between internal noise and memory, *Hearing Research* (2021), doi: https://doi.org/10.1016/j.heares.2021.108403

# Temporal integration for amplitude modulation in childhood:

# Interaction between internal noise and memory

**Laurianne Cabrera [1,2], Irene Lorenzini [1], Stuart Rosen [2],**

**Léo Varnet [3], Christian Lorenzi [3]**

[1] Université de Paris, INCC UMR 8002, CNRS, F-75006 Paris, France

[2] Speech, Hearing and Phonetic Sciences, UCL, UK

[3] Laboratoire des Systèmes Perceptifs (UMR 8248), CNRS, Ecole normale

supérieure, Université Paris Sciences & Lettres (PSL), Paris, France

*Corresponding Author*: Laurianne Cabrera, PhD

Integrative Neuroscience and Cognition Center (UMR 8002)

Université de Paris-CNRS

45 rue des saints pères

75006, Paris

laurianne.cabrera@parisdescartes.fr

Highlights:

- With increasing number of AM cycles, children and adults exhibit better AM detection thresholds
- 5-to 6-year-olds displayed worse thresholds with 2 AM cycles than other age groups
- Young adults showed better thresholds with 8 AM cycles than other age groups

- Changes in internal noise combined with memory capacities simulated those differences
- Thus, late processing stages for AM continue to develop late into childhood

**Abstract**

It is still unclear whether the gradual improvement in amplitude-modulation (AM) sensitivity typically found in children up to 10 years of age reflects an improvement in "processing efficiency" (the central ability to use information extracted by sensory mechanisms). This hypothesis was tested by evaluating temporal integration for AM, a capacity relying on memory and decision factors. This was achieved by measuring the effect of increasing the number of AM cycles (2 vs 8) on AM-detection thresholds for three groups of children aged from 5 to 11 years and a group of young adults. AM-detection thresholds were measured using a forced-choice procedure and sinusoidal AM (4 or 32 Hz rate) applied to a 1024-Hz pure-tone carrier. All age groups demonstrated temporal integration for AM at both rates; that is, significant improvements in AM sensitivity with a higher number of AM cycles. However, an effect of age is observed as both 5-6 year olds and adults exhibited more temporal integration compared to 7-8 and 10-11 year olds at both rates. This difference is due to: (i) the 5-6 year olds displaying the worst thresholds with 2 AM cycles, but similar thresholds with 8 cycles compared to the 7-8 and 10-11 year olds, and, (ii) adults showing the best thresholds with 8 AM cycles but similar thresholds with 2 cycles compared to the 7-8 and 10-11 year olds. Computational modelling indicated that higher levels of internal noise combined with poorer short-term memory capacities in children accounted for the developmental trends. Improvement in processing efficiency may therefore account for the development of AM detection in childhood.

## INTRODUCTION

Amplitude-modulation (AM) information represents the slow temporal variations in amplitude over time of incoming sounds. AM is essential for speech comprehension as slow modulations (< 5-8 Hz) convey important linguistic information (Drullman et al., 1994a, 1994b; Rosen, 1992; Varnet et al., 2017). To date, only a few studies have assessed in children the development of the ability to detect and discriminate slow AM fluctuations. Nevertheless, a typical development of slow AM perception is likely to be a pre-requisite of typical auditory and linguistic development (Goswami et al., 2002; Lorenzi et al., 2000; Witton et al., 2002).

Improvement in AM detection sensitivity during childhood may reflect the maturation of sensory-processing mechanisms leading to better temporal resolution with age, implying a better ability to follow fast changes in AM over time. Conversely, it is also possible that such improvement could reflect development in "processing efficiency". This would posit that the temporal resolution of the auditory system is adult-like but the ability to make optimal use of the extracted sensory information is not. This latter hypothesis was suggested in a pioneering study by Hall and Grose (1994). They showed that auditory sensitivity to sinusoidal AM, as measured by detection thresholds, improves with age between 5 and 10 years, but that AM sensitivity was similarly affected by AM rate (ranging from 5 to 200 Hz) at all ages. This suggests that sensory factors constraining temporal resolution are mature early on, as also suggested by studies with young infants (Walker et al., 2019). However, in a recent similar study, Buss et al. (2019) found evidence that temporal resolution may be poorer in children than in adults for low AM rates. In this study, 5-to-11-year-old children and adults had to detect a target sinusoidal AM applied to a 4300-Hz pure-tone carrier at three AM rates: 16, 64, and 256 Hz. The results replicated improvement in AM

detection thresholds with age at higher AM rates, 64 and 256 Hz, but this was less clear for the slower rate of 16 Hz. In another study, Peter et al. (2014) assessed AM detection thresholds in 8 to 11 year olds, 12- to 17 year olds and young adults using broadband noise carriers and AM rates ranging from 4 to 128 Hz. Although, they did not observe a significant interaction between age and modulation frequency, they observed slightly worse thresholds in 8- to 11-year-old children compared to older children for a 4 Hz modulation rate, but not at higher AM rates.

Note though that Hall and Grose (1994), Peter et al., (2014) and Walker et al. (2019) used a stochastic (broadband noise) carrier whereas Buss et al. (2019) used a deterministic one (a pure tone). It is thus possible that the external variability in acoustic stimuli may influence AM detection thresholds in childhood, and that temporal resolution even at low rates might improve between 5 and 11 years.

Another way to assess the development of sensory-processing mechanisms is to explore susceptibility to AM masking. It is now commonly assumed that AM perception reflects the operation of a bank of modulation filters selectively tuned for AM rate (Biberger and Ewert, 2016; Dau et al., 1997a) that are implemented centrally in the auditory system (Liégeois-Chauvel et al., 2004). "AM masking" effects are observed in psychophysical tasks using either two simultaneous sounds, a target and a masker, that fluctuate at a similar AM rate, or when the target and masking AM are simultaneously presented and applied to the same carrier. In such tasks, an elevation of detection thresholds for the target AM is observed, reflecting the selectivity of modulation filters for AM rate (Bacon and Grantham, 1989; Houtgast, 1989).

In an additional experiment, Buss et al. (2019) measured AM masking in children and adults. The results showed comparable effects of masker AM rate and a lack of interaction with age suggesting that modulation selectivity is mature by 5 years of age. Nevertheless, the overall effect of age on detection thresholds in the masked conditions indicates reduced efficiency in AM processing in childhood. Cabrera et al. (2019) explored further the development of AM masking in childhood and the role of "processing efficiency" (in the case of AM perception, the ability to make optimal use of the available temporal-modulation information at the output of modulation filters). It was hypothesized that poor AM detection thresholds in childhood may relate to higher levels of internal noise (neural variability at each level of the auditory system) and/or the use of sub-optimal decision strategies. In this study, detection thresholds were measured for slow sinusoidal AM (with a rate of 4, 8, or 32 Hz) applied to carriers whose inherent random modulations exerted different amounts of AM masking. More specifically three carriers were used: a pure tone, a narrowband noise with small inherent random AM fluctuations and a narrowband noise with higher inherent random AM fluctuations. Results showed that between 5 and 11 years, AM detection thresholds improved and surprisingly, that susceptibility to AM masking caused by the inherent random AM fluctuations of the narrowband noise carriers increased with age. Computer simulations of an auditory model based on the modulation-filterbank concept was used to simulate 1) poor sensory processing by changing the degree of selectivity of modulation filters; 2) poor processing efficiency by varying levels of internal noise at the output of the modulation filters; or 3) poor processing efficiency by simulating a suboptimal decision strategy. The model reducing internal noise levels with age by a factor 10 better accounted for the observed developmental trends

between 5 and 11 years. Altogether, the findings on AM masking suggest that at least some aspects of AM processing mature during childhood.

To explore further the development of processing efficiency for AM in childhood, the current study focused on a phenomenon known as "temporal integration" for AM cues. Several studies in adults have shown that AM detection improves with an increasing number of AM cycles (Dau et al., 1997b; Sheft and Yost, 1990; Viemeister, 1979; Wallaert et al., 2018, 2017, 2016). This so-called "temporal integration" is currently assumed to reflect high-level, i.e., central, decision processes involving "multiple looks" (Viemeister and Wakefield, 1991) or a template-matching process (Dau et al., 1997b). Consistent with this idea, electrophysiological work conducted on gerbils showed that the parietal cortex is required to temporally integrate AM information, and thus, accumulate sensory evidence for discrimination decisions. Yao et al. (2020) measured gerbils' discrimination abilities between 4 and 10 Hz AM modulated noises using a behavioural task. The stimulus duration varied from 100 to 2000 ms to assess the shortest stimulus duration for which animals could discriminate AM, reflecting integration time. Pharmacological inactivation of the parietal cortex was shown to increase minimum integration times. More specifically, when the excitatory projections from auditory cortex to parietal cortex were chemogenetically inactivated, the behavioural integration time was significantly reduced, showing the role of parietal cortex in temporal integration of AM cues.

A recent investigation of temporal integration conducted with adult human listeners combining a psychophysical and modelling approach indicates that for low AM rates (below 5 Hz), temporal integration is also constrained by the limited capacity of short-term auditory memory (Wallaert et al., 2017, 2018). In these studies, AM detection thresholds are measured in adult listeners as a function of the number of AM cycles

7

available in the stimuli. A computational model was used to simulate the thresholds implementing different sources of internal noise supposed to limit the performance when processing AM. An additive noise was introduced to limit intensity discrimination and AM sensitivity. An additive time-varying memory noise was introduced to simulate imperfect retention of temporal-envelope information in each observation interval. All internal noises were modeled as Gaussian noises, and added independently to the output of the bandpass AM filter in the model.

In children, previous studies used only fixed stimulus durations to assess the effect of different AM rates on detection thresholds. Peter et al. (2014) suggested that the slightly worse thresholds observed at low AM rates may relate to the low number of cycles available in the target signal when the stimulus duration is kept constant at 500 ms. In their study, the 4-Hz rate involved the lowest number of AM cycles, 2 cycles, compared to the other AM rates tested. In contrast, Lorenzi et al. (2000) varied the target-stimulus duration from 500 to 1500 ms for a 4 Hz modulation rate and observed similar temporal integration between adults and six children between 8 and 15 years, suggesting no improvement between childhood and adulthood.

The present study sought to measure temporal integration at two different AM rates in a large cohort of children aged between 5 and 11 years, in order to probe the integrity and maturation of memory and decision processes in the AM domain. The main goal of the present study was to: 1) evaluate the effect of increasing the number of AM cycles (2 vs 8 cycles) on detection thresholds between 5 and 11 years for two slow AM rates (4 and 32 Hz) and, 2) use computational modelling to test whether AM-detection data in children were better simulated by changes in the characteristics of

AM filtering, internal noise and/or short-term memory capacities. It was assumed that temporal integration for AM detection should reflect the operation of the late, decision stage of AM processing (Dau et al., 1997b). It was also assumed that at the lower AM rate (4 Hz) where stimuli are longer in duration, temporal integration should reflect the additional constraints imposed by the limited echoic memory buffer involved in AM processing (King et al., 2019; Wallaert et al., 2018, 2017). If all age groups are similarly affected by the number of AM cycles or by AM rates, this would indicate that temporal integration is developed by 5 years and that central aspects of processing efficiency related to short-term memory and decision making for AM processing are well developed in childhood. However, if younger and older children are affected differently by the number of AM cycles and by AM rates, this would suggest that memory and decision mechanisms are still evolving and thus, that processing efficiency for AM continues to develop into childhood.

**METHODS**

**Participants**

Seventy-two children and 30 adults were included in this experiment (see also Cabrera et al., 2019). All participants reported typical cognitive development. Consent was obtained from parents and adult participants as approved by the university ethics committee. Adult participants received a monetary compensation for their time and children collected stickers on a science certificate to keep them motivated.

For both children and adults, absolute thresholds were assessed for both ears at octave frequencies between 0.25 and 8 kHz prior to testing. Two participants (one child and one adult) were excluded because their absolute thresholds for pure tones were above the normal range (>20 dB HL).

An additional 9 children were tested but not included in the final sample: three were at floor levels of performance for more than the half of the conditions (all aged 5-6 years), one withdrew from the study, four were not at school during testing, and one was excluded because of experimental error. Four adults were not included in the final sample because of missing data due to experimental errors. Furthermore, an outlier labelling rule (Hoaglin and Iglewicz, 1987) was applied to the data set. A total of 10 outliers were flagged and their data were not included in the analyses (see Cabrera et al., 2019).  Four outliers were flagged in the 5–6 year-old group, two in the 7–8 year olds, four in the 10–11 year olds, and two in the adult group (11% of the data). When those outliers were removed, the data did not differ from normality (Kolmogorov-Smirnov test, all ps > 0.05).

The final sample included: 21 5-6-year-olds (10 females; mean age = 5.7 years, SD = 0.4), 27 7-8-year-olds (13 females; mean age = 7.8 years, SD = 0.5), 24 10-11-year-olds (12 females; mean age = 10.7, SD = 0.4) and 30 adults (22 females, mean age = 22.5 years, SD = 2.5).

**Stimuli**

A full factorial design led to four experimental conditions:  two target AM rates ($f_m$: 4 and 32 Hz) presented for two numbers of modulation cycles (2 and 8). As previous studies showed that optimal AM detection is observed with 4-5 cycles, we selected 8 cycles to ensure that listeners are presented with a sufficient number of cycles (Edwards et al., 2008; Lee and Bacon, 1997). In all conditions, the stimuli included 50-ms raised-cosine onset/offset ramps, and the inter-stimulus interval was 500 ms. Standard sounds were not modulated in amplitude, and target sounds were modulated at depths ranging from $m$ = 100% to $m$ =1 %, in 20 steps of 2 dB. The starting phase

of the modulation was randomized on each trial. The carriers were sine tones centered

at 1027 Hz, generated with a random starting phase at a sampling frequency of 44.1

kHz. **Figure 1** shows the waveforms played in the 4 conditions.
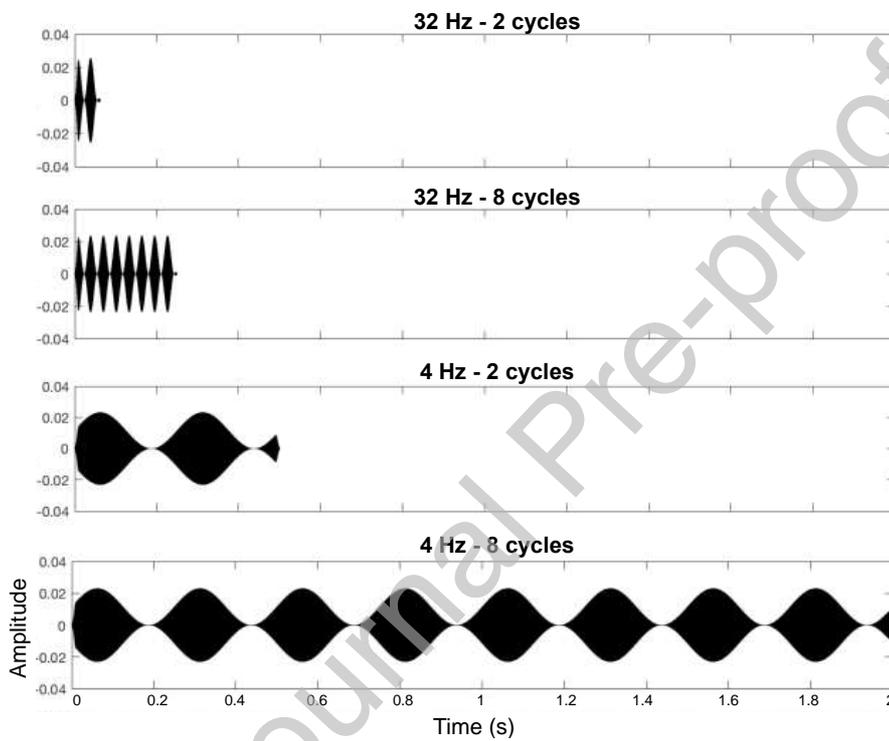
**Figure 1.** Waveforms of one stimulus exemplar in each condition. The target sounds modulated at 32 Hz were 62.5- and 250-ms long in the 2 and 8-cycle conditions, respectively (first two upper rows). The target sounds modulated at 4 Hz were 500- and 2000-ms long in the 2 and 8-cycle conditions, respectively (first two lower rows).

12

**Procedure**

The present experiment was part of a larger project on temporal processing capacity in childhood assessing modulation masking as well (see Cabrera et al., 2019). Children completed the whole experiment within 4 sessions of 20 min over a 2-week period and adults in one session of 1.5 hr. The 4 experimental conditions considered here occurred randomly during the whole experiment. AM detection thresholds were measured using a three-interval, three-alternative forced-choice (3I-3AFC) adaptive procedure, implemented on a touch-screen tablet. The sounds were played diotically through headphones (Sennheiser HD 25-SP II) at a level of 65 dB SPL.

In this task, participants were asked to listen to three animal characters who would appear on the screen and produce a sound one after the other and to find the odd-one out, i.e., the modulated target. On each trial, participants had unlimited time to respond and received visual feedback as to the correctness of their response. The next trial started automatically 600 ms after the participant's response.

The response on the first trial was always ignored and a one-down, one-up rule was used until the first reversal (Baker and Rosen, 2001). Then, a 2-down 1-up adaptive procedure was used to track the 71 % correct-point (Levitt, 1971). The first trial started at $m$ = 100 % and the initial step-size (the change in modulation depth) was 6 dB. This step size was reduced to 4 dB after the first reversal, and to 2 dB after the second reversal, the minimum $m$ was 1 %. The run stopped after the 8th reversal or after 32 trials. One estimate was collected for each participant in each condition and a second estimate was collected if fewer than 5 reversals were obtained, or if the track did not converge. The threshold in dB was the geometric mean of the last four reversals.

# RESULTS

## AM detection thresholds (AMDTs)

**Figure 2** represents the AM-detection thresholds (AMDTs) according to the number of AM cycles and modulation rate. All age groups show lower (better) AM detection thresholds in the 8-cycle conditions compared to 2-cycle conditions.

A mixed ANOVA assessed the effect of the within-subject factors of Cycle (2 levels: 2 and 8 cycles) and Rate (2 levels: 4 and 32 Hz) and the between-subjects factor of Age (4 levels: 5-6 years, 7-8 years, 10-11 years and adults) on the AMDTs. The analysis revealed a significant main effect of Rate [$F(1,98) = 13.34$, $p < .001$, $\eta^2 = .12$], Cycle [$F(1,98) = 518.16$, $p < .001$, $\eta^2 = .84$], Age [$F(3,98) = 13.13$, $p < .001$, $\eta^2 = .29$] and a significant interaction between Cycle and Age [$F(2,98) = 5.60$, $p = .001$, $\eta^2 = .15$]. Other comparisons were not significant [Rate x Age $F(3,98) = .34$, $p = .798$; Rate x Cycle $F(1,98) = 1.03$, $p = .312$; Rate x Cycle x Age $F(3,98) = .43$, $p = .735$].

For the main Age effect, pairwise comparisons with Bonferroni corrections showed that, overall, 5-6 years have higher (worse) AMDTs than 7-8,10-11 years and adults ($ps < .01$). The groups of 7-8 years showed similar AMDTs than 10-11 years ($p > .05$) but worse AMDTs than adults ($p = .016$). The group of 10-11 years showed overall similar AMDT than the group of adults ($p = .11$). Moreover, AMDTs were overall lower (better) at a high rate than at a low rate.

The analyses of the Cycle and Age interaction indicated that all age groups show temporal integration, that is, lower (better) AMDTs with increasing number of AM cycles ($ps < .001$). It is important to note that 5-6-year-olds showed worse AMDTs in the 2-cycle condition compared to the three other age groups ($p < .001$), and that

adults showed significant better AMDTs in the 8-cycle condition compared to all the

other age groups (p < .001).                                                          15
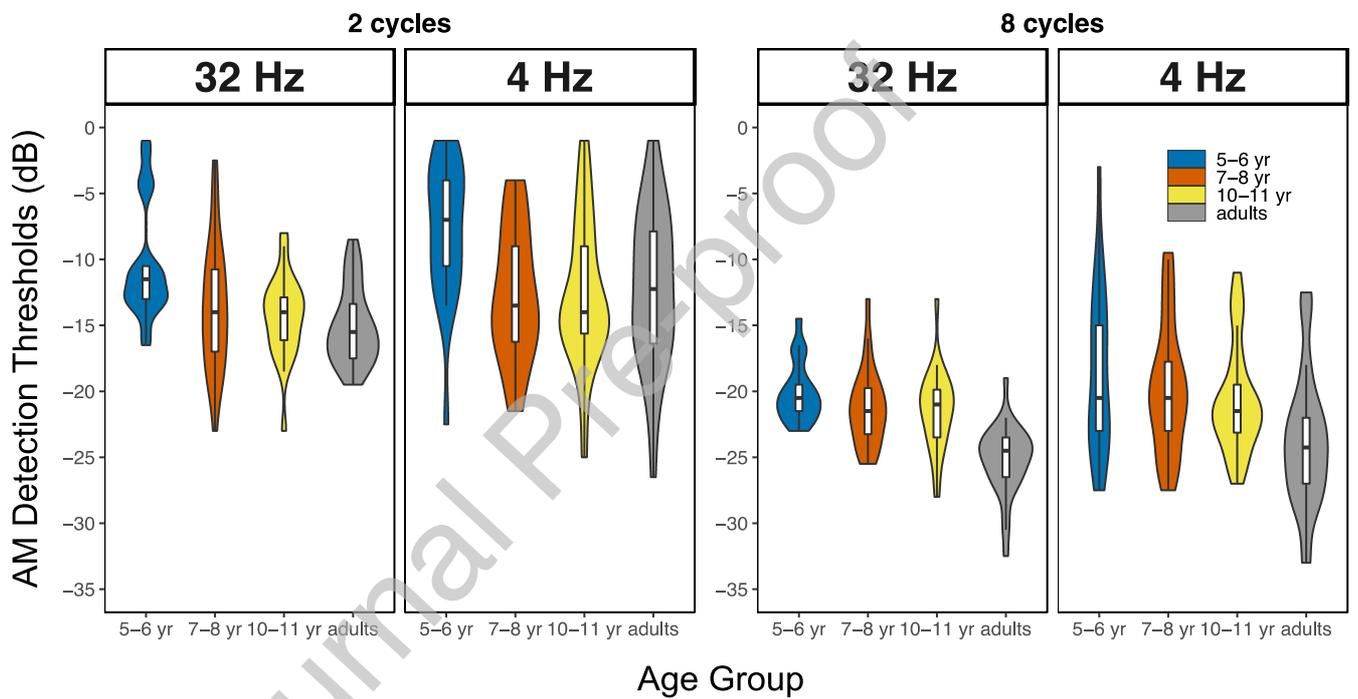
**Figure 2.** AMDTs (in dB) for each age group. The left and right panels represent thresholds in the 2-cycle and 8-cycle conditions, respectively, at each modulation rate (4 and 32 Hz). The boxes in the boxplots represent the 25th, 50th and 75th percentile for each

group/condition, with whiskers extending to the full range of the data. The violin plots illustrate the kernel probability density (i.e. the width of the violin area represents the proportion of the data located there).

**Temporal-integration effect**

A second ANOVA was run on the temporal integration scores shown in **Figure 3,** that are estimated as the difference in AMDTs between 2 and 8 cycle conditions, with Rate (2 levels) as a within-subject factor and Age (4 levels) a between-subject factor. This score represents the improvement in AMDTs when increasing the number of AM cycles from 2 to 8. The analysis revealed no significant effect of Rate [$F(1,98)$ = 1.03, $p$ = .312], nor significant interaction between Rate and Age [$F(2,98)$ = 0.74, $p$ = .735], but a significant effect of Age [$F(3,98)$ = 5.60, $p$ = .001, $\eta^2$ = .15]. Further analyses of the Age effect showed that 5-6 years showed slightly *more* temporal integration than the 7-8 years and 10-11 years ($ps$ = .085 and .16 with Bonferroni corrections, respectively), but showed a similar level of integration to adults ($p$ = 1). Interestingly, adults showed higher integration than the 7-8 years and 10-11 years ($p$ = .007 and .015, respectively). This is related to the worse AMDTs shown by the youngest group in the 2-cycle condition and to the best AMDTs observed in the 8-cycle condition in adults, making these two groups exhibiting a larger difference between the 2 and 8 cycle conditions.

To sum up, all child groups showed significant effects of AM rate and AM cycle, but 5-6 years and adults exhibited similar levels of temporal integration for AM detection. Indeed, at both rates, the youngest age group showed significantly higher (worse) AMDTs in the 2-cycle condition and reached similar thresholds in the 8-cycle condition compared to the groups of 7-8 years and 10-11 years. Furthermore, adults outperformed all child groups in the 8-cycle conditions indicating that temporal integration continues to develop after 11 years of age.
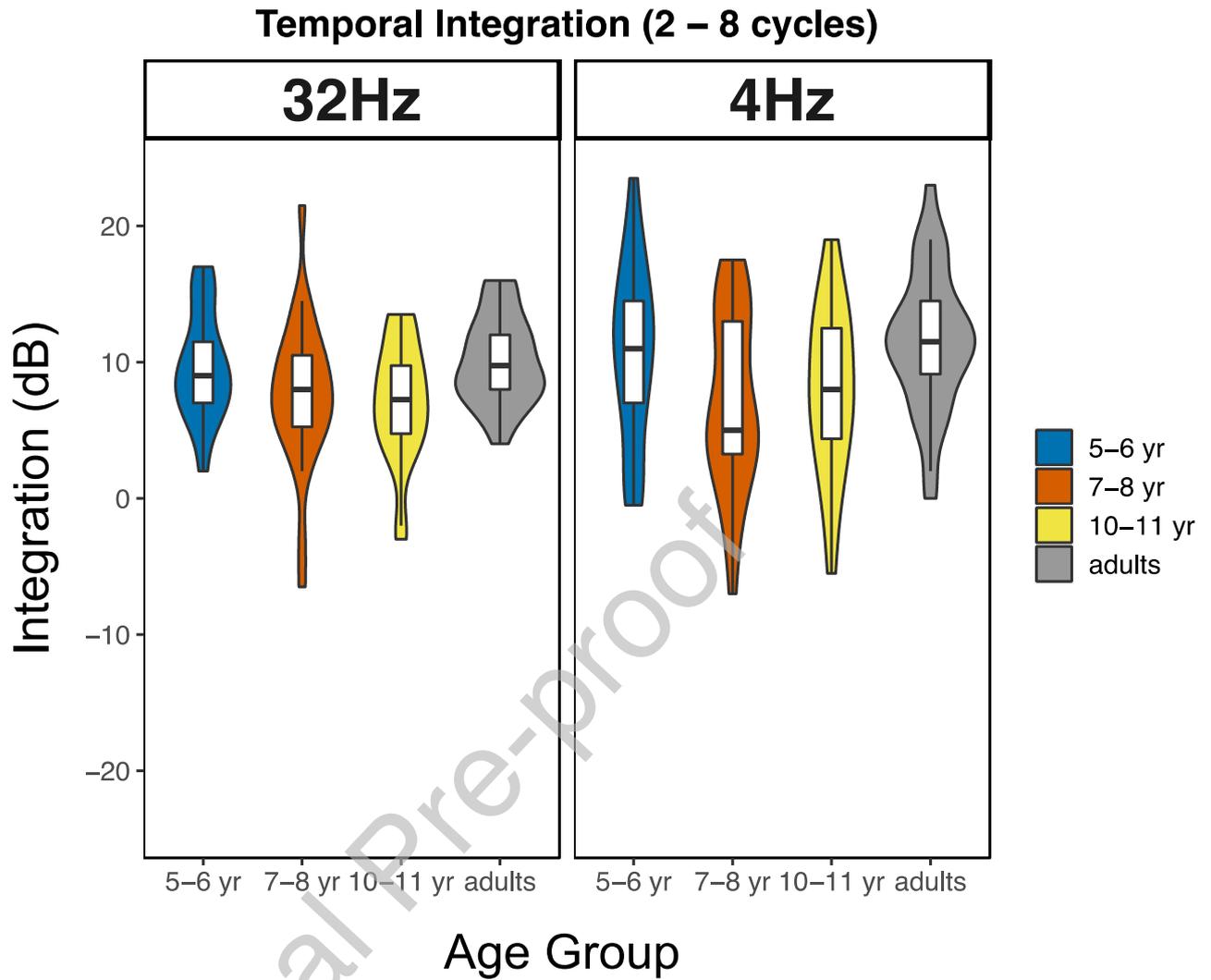
**Temporal Integration (2 – 8 cycles)**



**Figure 3.** Integration scores (in dB) for each age group. Box- and violin plots are as described for Figure 2.

1   **Modelling study**

2       A computational model of AM processing was used to assess whether the

3   differences between age groups could relate to changes in (low-level) sensory

4   processing or in (high-level) processing efficiency. The same model as in Cabrera et

5   al. (2019) was used to simulate AMDTs while manipulating parameters at different

6   stages of AM processing. Three main processing stages were represented in this

7   model: *Early* (i.e., peripheral) processing (bandpass cochlear filtering, instantaneous

8   amplitude compression, half-wave rectification, short-term adaptation), *mid-level*

9   processing (envelope extraction via a bank of selectively-tuned AM filters with a

10  constant Q value of 1), and *late* processing (additive internal noise; exponential

11  memory noise with a 1.4-sec time constant, template-matching decision device). The

12  model structure is detailed in the Appendix. The "late processing" stage implements

13  central factors constraining processing efficiency, that is, the ability to make optimal

14  use of the available AM information extracted by sensory mechanisms at early and

15  mid-level stages. This model previously suggested that AM-masking data in childhood

16  were better predicted by a drastic change in the additive internal noise between 5 and

17  11 years (by a factor 10).

18

19      *Effects of varying the magnitude of the additive internal noise*

20      In the current study, the first model, named *Model A*, uses the standard

21  deviation (SD, in model units, m.u.) of the additive internal noise that best fitted the

22  AMDT data of Cabrera et al. (2019). More precisely, Model A uses the SD of the

23  internal noise that maximized the model's "goodness-of-fit" at the 8 Hz AM rate. This

24  was achieved by minimizing the model's mean absolute error, evaluated as the

25  average absolute value of the difference between simulated AMDTs and "real" AMDTs

1 averaged across participants for each age group in Cabrera et al. (2019). The SD of

2 the internal noise is set to 2 x $10^{-4}$ for adults, 5 x $10^{-4}$ for 10–11-year-olds, 1 x $10^{-3}$, for

3 7–8-year-olds, and 2 x $10^{-3}$ for 5–6- year-olds. So, the SD of internal noise varies by

4 a factor 10 between 5-6-year-olds and adults. Real averaged data and simulated data

5 are shown in **Figure 4**. As expected, Model A yielded better AMDTs with decreasing

6 level of internal noise. At both AM rates, the improvement in AM sensitivity is larger

7 for 8 AM cycles (∼5 dB) than for 2 AM cycles (∼2 dB), predicting increased temporal

8 integration over age. However, Model A does not fit the real data well, specially for the

9 2-cycle and 32-Hz condition, with the model prediction error averaged across age

10 groups = 4.7 dB.

11 Model A reproduces well the integration data for the group of 5-6 years at 32

12 Hz and 4 Hz but not for the other groups of children nor for the adult group specially

13 at 32 Hz. According to this model, a decrease in the level of the additive internal noise

14 by a factor of 10 is associated with an *increase* in temporal integration by about 4 dB

15 at 4 Hz and 7 dB at 32 Hz. In contrast with the simulated data, temporal integration

16 *decreases* in real listeners for intermediate age groups (7-6 years and 10-11 years).

17 The failure of Model A at 32 Hz relates partially to the short duration of the stimuli

18 when using 2 AM cycles (62.5 ms). In this case, the response of the simulated

19 modulation filter tuned to 32 Hz (target rate) is dominated by the response of the filter

20 to the stimulus onset (see Edwards et al., 2008; Sheft and Yost, 1990). The latter is

21 superimposed on the target and masks its temporal structure, resulting in poor

22 detection of the AM fluctuation. Thus, the following paragraphs explore the role of two

23 specific parameters related to the model's goodness of fit only for the 4-Hz conditions.
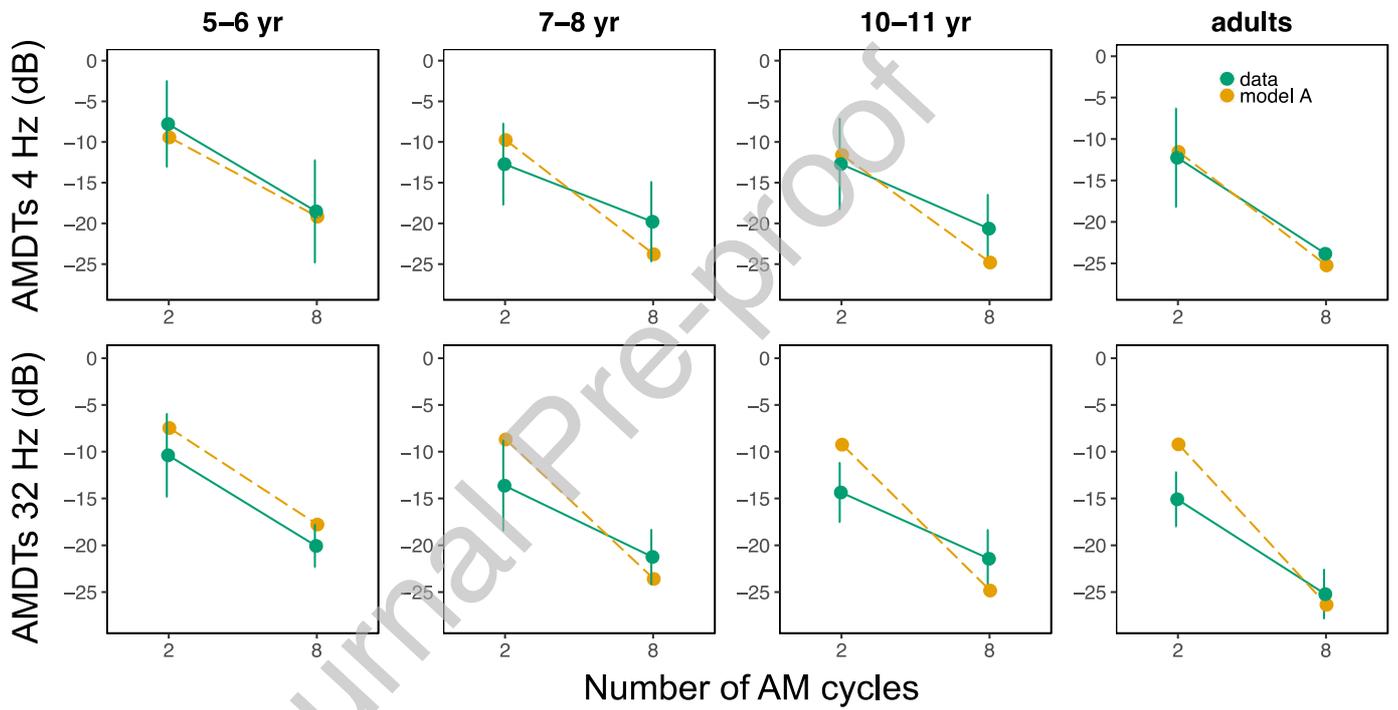
**Figure 4.** Temporal integration effects: AMDTs (in dB) for each age group (from left to right column: 5-6 years, 7-8 years, 10-11 years and adults) as a function of number of AM cycles (2 vs 8 cycles) at each modulation rate (4 Hz: upper row; 32 Hz: lower row). The

22

1    real data are represented with the green circles and solid lines. The simulated data with Model A are represented with the orange

2    circles and dashed lines.

1    *Effects of varying the strength of short-term adaptation*

2    At the early and mid-level processing stages, two crucial aspects of temporal

3    processing can be modified in the model: the highpass filtering of the compressed

4    output of cochlear filters which simulates short-term adaptation effects in the temporal-

5    envelope domain and AM selectivity as reflected in the bandwidths of the modulation

6    filters. The effect of changing the latter parameter was tested in Cabrera et al. (2019).

7    A progressive increase in AM selectivity (in other words, a reduction in the bandwidth

8    of modulation filters) as a function of age did not account for real AM detection data.

9    The effect of changing the short-term adaptation parameter is tested here.

10   In Model A, the cutoff frequency of the highpass filter (rolling off at 20 dB per

11   decade) applied to the output of all peripheral (cochlear) channels is initially set to 2.5

12   Hz as in Wallaert et al. (2018; see also Tchorz and Kollmeier, 1999) to simulate short-

13   term adaptation effects observed in adult listeners. The cutoff frequency of the

14   highpass filter is modified in *Model B*. Increasing the cutoff frequency of the highpass

15   filter has the effect of enhancing the transient response of the model at the stimulus

16   onset. Here, it is hypothesized that the youngest children may be more distracted by

17   this transient signal than older children. For each AM rate, this "bottom-up" attentional

18   effect would result in a greater detrimental effect on AM detection at the shortest

19   stimulus duration (i.e., 2 cycles), increasing consequently temporal integration for AM

20   detection. Although the current model was not designed to account for sustained

21   attentional effects in AM perception, its decision stage based on a template-matching

22   strategy (see Appendix section) could nevertheless be influenced by the presence of

23   strong onset cues at the output of the simulated cochlear filters. The cutoff frequency

24   of the highpass filter was increased from 2.5 Hz (Model A) to 3.5 Hz (Model B1) or

25   decreased to 1.5 Hz (Model B2). As observed in **Figure 5.a** and **5.b**, increasing the

1    cutoff frequency of the highpass filter to 3.5 Hz (Model B1) has the effect of elevating

2    AMDTs for the 4 Hz-2 cycle condition for the groups 7-8 years and 10-11 years, by 1

3    and 2 dB, respectively, making model predictions more consistent with the data (model

4    fitting error < 1 dB, **Figure 5.b**). However, for the 4 Hz-8 cycle condition, changing the

5    cutoff frequency of the highpass filter does not improve the fit for any group. **Figure**

6    **5.c** shows the effect of changing the adaptation parameter at the output of the

7    adaptation filter of the model that influences slightly the transient response over the

8    first 0.2 sec of the stimulation. In this case, the model's fitting errors for the integration

9    effect are still greater than 4 dB for those two groups. Manipulating the characteristics

10   of the adaptation stage (data not shown) has little effect on the predicted AMDTs for

11   the 32-Hz AM rate. These simulations suggest that children do not place more weight

12   on the onset of the signal or that this effect varies over age. Overall, the model using

13   a cutoff frequency of 2.5 Hz (Model A) gives a better account of real data for all groups,

14   suggesting that the effects of short-term adaptation (Models B1 and B2) do not differ

15   markedly over age.

16

a
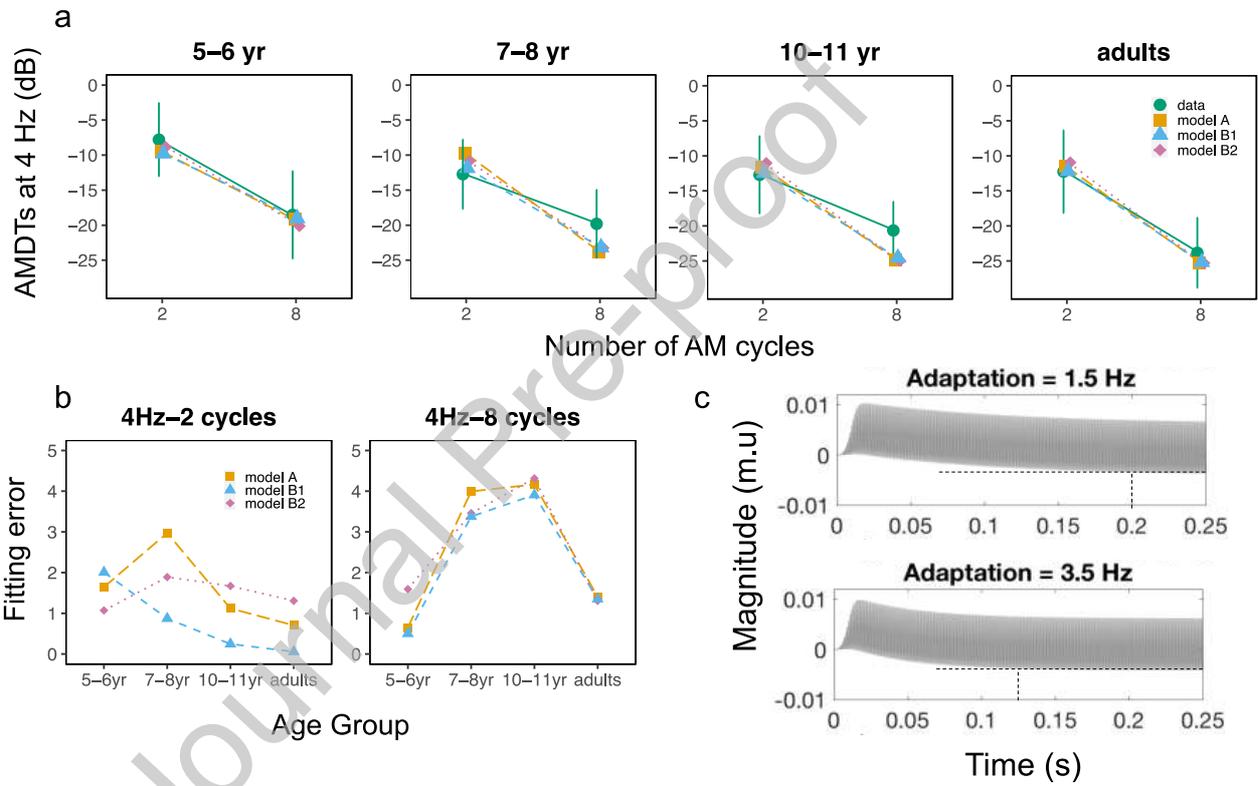
1

**Figure 5. a.** AMDTs (in dB) at 4 Hz as a function of number of AM cycles for each age group. Each panel shows real data (green circles) and simulated data (Model A: orange squares; Model B1: blue triangles; Model B2: pink diamonds). Simulations assess changes in internal noise only (Model A), and changes in internal noise and adaptation [Model B1 with stronger adaptation (cutoff frequency of highpass filter = 3.5 Hz) vs Model B2 with lower adaptation (cutoff frequency of highpass filter=1.5 Hz)]. **b.** Model's goodness-of-fit, estimated by the fitting error (absolute difference between observed and simulated data) of each model with 2 cycles (left panel) and 8 cycles (right panel) as a function of age group. **c**. Output of the adaptation stage of the model, for a cutoff frequency of the highpass filter set at 1.5 Hz (top panel) and 3.5 Hz (lower panel). Note that the transient response is stronger (faster decay) for a 3.5-Hz cutoff frequency.

1

2          *Effects of simulating imperfect retention of temporal-envelope information*

3          As in Wallaert et al. (2017), we simulated poorer processing efficiency for adult

4   listeners by including an additive time-varying memory noise at the output of

5   modulation filters. The standard deviation of this memory noise increased backward

6   in time according to an exponential function using a 1.2-sec half life. Memory noise

7   aimed to simulate imperfect retention of temporal-envelope information in each

8   observation interval. This is especially important for long stimuli, and thus, for the 4

9   Hz-8 cycles condition where the stimulus duration is equal to 2 sec. **Figure 6.a** shows

10  the effect of introducing this exponential memory noise on the predicted AMDTs for

11  the 4-Hz condition using two different SD values for the memory noise (*Model C1*: 3e-

12  4 m.u., and *Model C2*: 5e-4 m.u.) but the same half life. It is also important to keep in

13  mind that Models C1 and C2 also use the levels of additive internal noise of Model A.
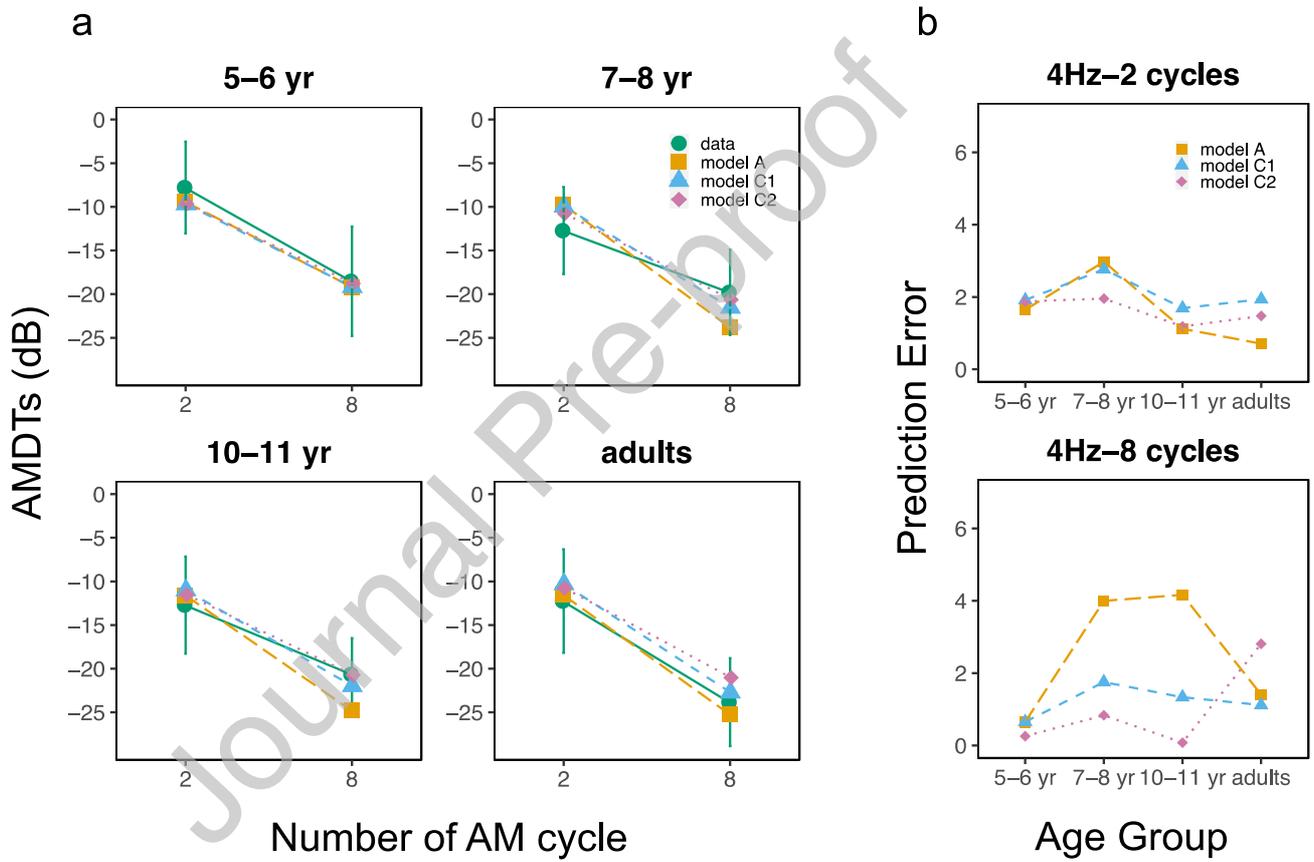
14         Increasing the level of memory noise does not improve the model's goodness-

15  of-fit for the group of 5-6 years and has little effect for the adults as observed in **Figure**

16  **6.b**. However, manipulating this parameter in the model influences the goodness-of-

17  fit for the two other age groups and especially with 8 AM cycles. The effects of the

18  memory noise on the internal representation of temporal envelopes are illustrated in

19  **Figure 7** for a 4-Hz AM stimulus. Overall, the introduction of memory noise improves

20  the model's goodness-of-fit of AMDTs for the 4-Hz condition between 7 and 11 years.

21  As expected, this improvement is more visible for the longest stimuli, i.e., 8 cycles.

22  This suggests that the fidelity of information retention and the size of the short-term

23  memory buffer influence the model's behavior for the two older child groups. The

24  higher the variance of memory noise, the worse the fidelity of the former memory trace

25  of the temporal envelope. The level of memory noise has little effect on the fit of the

1  model to the results from the youngest children, because they have greater internal

2  noise. Using the higher level of memory noise improves the fit of the model for the

3  older children because the level of the internal noise is lower in these groups. For

4  adults, Model A without additional memory noise provides the best fit of the data.

5  Therefore, memory noise for AM processing seems to decrease over age, and

6  importantly, needs to be appreciated in relation with a decrease in internal noise during

7  childhood.

1

a

### 5–6 yr

### 7–8 yr

- ● data
- ■ model A
- ▲ model C1
- ◆ model C2

### 10–11 yr

### adults

AMDTs (dB)

Number of AM cycle

b

### 4Hz–2 cycles

- ■ model A
- ▲ model C1
- ◆ model C2

Prediction Error

### 4Hz–8 cycles

Age Group

30

1    **Figure 6. a.** AMDTs (in dB) at 4 Hz for each age group as a function of the number of AM cycles (2 vs 8 cycles) for the real

2    data (solid lines) and the simulated data according to changes in internal noise only (Model A, orange squares and long dashed

3    lines), changes in internal noise and memory noise (Model C1 with low memory noise, blue triangles and dashed lines vs Model C2

4    with higher memory noise, pink diamonds with dotted lines). **b.** Model's fitting errors (absolute difference between real and simulated

5    data) of each model with 2 cycles (left panel) and 8 cycles (right panel) as a function of age group.

6

7

8

9

10

11

12

13
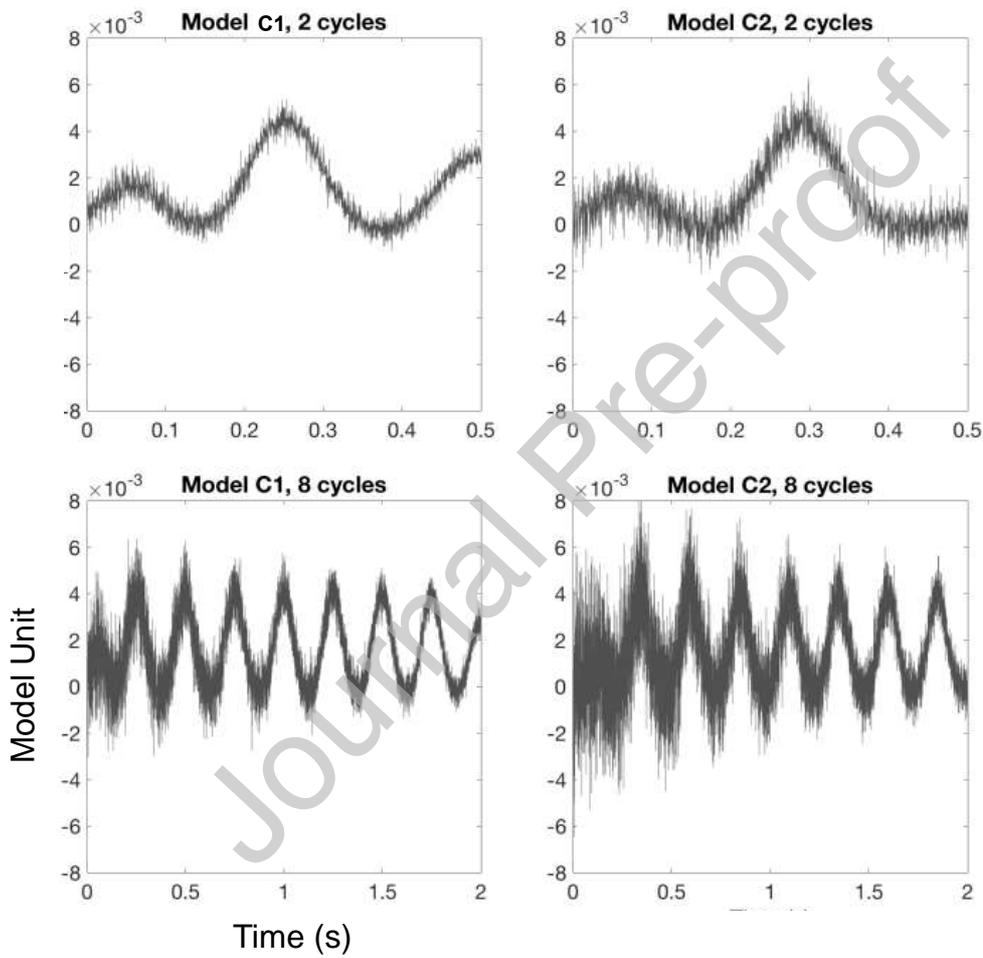
14

15

16

31

1



**Figure 7**. Internal representation of the AM stimulus computed by the model for two levels of memory noise (low: Model C1, and high: Model C2) in the conditions 4 Hz-2 cycles and 4Hz-8 cycles.

32

1    **DISCUSSION**

2       The present study reveals that children spanning the age of 5 to 11 years

3    demonstrate temporal integration for AM detection. All benefit from an increased

4    number of AM cycles when detecting slow and fast AM, and thus demonstrate a

5    capacity to accumulate sensory evidence as to the presence of an AM signal.

6    Surprisingly, greater temporal integration is observed for the youngest children, 5-6

7    years, compared to the older children, 7-8 years and 10-11 years. This effect results

8    from higher (worse) detection thresholds with 2 AM cycles at 5-6 years but similar

9    thresholds with 8 AM cycles compared to the older children. Overall those results

10   suggest that sensory-processing mechanisms related to temporal integration of AM

11   information are mature by 5 years, but that higher stages of AM processing are still

12   developing.

13      Importantly, no effect of age between 5 and 11 years is observed with a higher

14   number of AM cycles. The 8-cycle condition represents the only condition where

15   younger children show AM detection thresholds similar to those of older children (see

16   Cabrera et al., 2019). This result suggests that the poorer thresholds usually observed

17   in younger children for AM detection tasks are not due to an overall poorer ability in

18   performing such a psychophysical task. The fact that this effect is observed at two AM

19   rates also suggests that children's responses are independent of the absolute duration

20   of the sounds. If their poorer thresholds were related only to a lack of attention, we

21   should have observed poor thresholds in the shortest-duration condition only, that is,

22   at 32 Hz and with 2 AM cycles (duration ~62.5 ms), and not at 4 Hz and with 2 cycles

23   (duration ~250 ms). The higher (worse) thresholds observed for young children may

24   then be related to their difficulty in processing AM cues when the modulated target

25   sound shows a small number of AM cycles (as suggested by Peter et al., 2014). This

1 specific result, unrelated to sound duration, may relate to *inefficiencies in processing*

2 *the temporal-envelope information when few cycles are available.*

3 To explore further the stages of AM processing that may develop over age, we

4 used a computational model based on the modulation-filter bank concept simulating

5 peripheral and central constraints in AM processing. The results of the modelling study

6 reveal that, contrary to a previous study on AM masking (Cabrera et al., 2019), the

7 observed data are not well simulated by a simple decrease in the level of internal noise

8 over age: A progressive decrease in the level of internal noise during childhood

9 predicts better temporal integration with increasing age while the real data show better

10 temporal integration in the youngest group (5-6 years). Additional simulations indicate

11 that changes in short-term adaptation cannot account for the change in temporal

12 integration over age; changing the weight of transients does not influence the model's

13 predictions for 5-6-years or adults, nor does it improve the model's predictions for the

14 other two child groups. Changing the characteristics of adaptation provides a slightly

15 better fit for the 2-cycles condition at 7-8 years but this improvement is not sufficiently

16 large. Thus, children do not seem to weight differently transient responses over age.

17 Finally, the only aspect of AM processing changing over age that yields better

18 predictions for the 7-8 and 10-11 year groups is imperfect retention of temporal-

19 envelope information, as proposed by Wallaert et al. (2017). In the present model, this

20 constraint is simulated by degrading the temporal-envelopes of incoming signals at

21 the output of modulation filters with an additive, exponential memory noise. Temporal-

22 envelope cues are more masked by memory noise near the onset of the stimulus.

23 Increasing the SD of the memory noise lengthens the initial portion of the stimulus that

24 is most impacted by the memory noise, reducing by the same amount the length of

25 the final portion that makes an effective contribution to decision making. The imperfect

1    retention of temporal-envelope information in each observation interval necessarily

2    has a greater detrimental effect on the detection of the longest stimuli, e.g., at 8 AM

3    cycles. The fact that increasing memory noise only improves model predictions for the

4    7-8 years and 10-11 years suggests that: 1) The level of the additive internal noise at

5    5-6 years is high enough to dominate any effect of memory noise; 2) the level of the

6    memory noise decreases over age, as adults showed significantly better thresholds

7    than the child groups in the 8-cycle condition; 3) AM processing is still not fully mature

8    in late childhood, but this may relate to inefficiency at higher levels of processing, i.e.,

9    echoic memory.

10

11       Echoic memory, or 'auditory sensory memory', is the ability to retain a sound

12    stimulus right after its occurrence, allowing further processing. This type of memory is

13    described as a pre-attentive phase of 100-300 ms where auditory information is

14    temporarily stored for further manipulation (Cowan, 1984; Massaro, 1972). So far,

15    developmental studies have shown that this cognitive function improves throughout

16    childhood and reaches its peak in early adulthood. Evoked-potential investigations

17    have tested the short-term retention and processing of tones, in children *vs* adults,

18    through oddball paradigms assessing the Mismatch Negativity (MMN) to deviant tones

19    differing in frequency or duration. These studies have globally revealed reduced

20    amplitude and higher latency of the MMN up to 10 years of age, which is thought to

21    reflect the shorter duration and faster decay of the sensory memory trace (see Bartha-

22    Doering et al., 2015 for a review). Keller and Cowan (1994) reported comparable

23    results from a 2I-2AFC tone-comparison task, and concluded that the persistence of

24    echoic memory for tones is shorter in children of 6-to-7-year-olds than adults. The

25    clear-cut differences in echoic memory performance between children and adults may

1    explain why the model predictions for AM detection are improved by simulating

2    imperfect retention of temporal-envelope information in childhood.

3

4    The dynamics behind the protracted improvement of echoic memory are not yet clear.

5    It has been proposed that a general increase in global processing speed might

6    mediate this process (Ferguson and Bowey, 2005). Neural maturation of specific brain

7    areas may also play a role in improvement of echoic memory. Recent

8    neurophysiological studies in animals suggest that parietal cortex activity relates to

9    the integration of AM information (Yao et al., 2020). When excitatory auditory cortex

10   inputs to parietal cortex are inhibited, this results in reduced temporal integration in

11   gerbils for AM-target stimulations of 100, 300 and 600 ms. When parietal cortex is

12   directly deactivated, increased errors at long stimulus durations, of 1 or 2 sec, are

13   observed, suggesting that the parietal cortex integrates sensory input from the primary

14   auditory cortex and thus, plays an essential role in an auditory decision task.

15          Whether parietal cortex is more specifically involved in information retention is

16   still under investigation. Lesion studies suggest that parietal areas are involved in

17   memory judgments through decision making. For instance, patients with parietal

18   lesions show more difficulties in integrating new external cues with internal memory

19   evidence (Dobbins et al., 2012). Finally, the role of prefrontal areas in decision making

20   (Hanks et al., 2015), as well as of auditory areas in auditory short term memory (Scott

21   et al., 2014) need to be specified in such temporal integration tasks. Better knowledge

22   of the specific development of those cortical regions and their implications for AM

23   processing would help to better understand what stages of processing are still

24   developing through childhood. Computational modelling may also help in future

1  studies to simulate the development of memory processing and of decision making in

2  AM detection.

3      It is worth noting that our current model did not take into account the potential

4  detrimental effects of memory capacities on the whole sequence of stimuli composing

5  a given trial in the 3AFC procedure. Memory limitations may have not only affected

6  the internal representation (and especially the onset) of each stimulus, but also the

7  whole sequence of stimuli composing a given trial. This may especially be the case

8  when using 8 cycles at 4 Hz, where the whole trial lasts for 7 s. In other words, the

9  internal representation of the first stimulus of a given trial might have been more

10  affected by memory limitations than the internal representation of the second or last

11  stimulus. The current modelling architecture aimed to reproduce aspects of echoic

12  memory in the AM domain, not such effects over the whole trial. For 7-s long

13  sequences of sounds, other aspects of memory may come into play such as working

14  memory. Further work is warranted to investigate such constraints in temporal

15  envelope retention. Still, using a 3AFC procedure, where the position of the target trial

16  is randomized on each trial, this effect should be averaged out in the results. It is also

17  noteworthy that the current model succeeded in reproducing the 4-Hz data without

18  taking account of the temporal structure of the whole trial.

19      Finally, our model did not take attentional effects into consideration. Wright and

20  Dai (1998) showed that expectations influence AM detection performance (see also

21  Conroy and Kidd, 2021). To the best of our knowledge, there is no model of auditory

22  modulation processing able to reproduce such uncertainty effects.

23

24  **CONCLUSIONS**

1    The present study indicated that children from 5 to 11 years, as well as adults,

2    showed an improvement in AM detection thresholds when presented with more AM

3    cycles in the modulated target. Nevertheless, this effect is influenced by age, as the

4    youngest children of 5-6 years displayed the worst thresholds when only 2 cycles were

5    available, irrespective of the AM rate to detect. Furthermore, adults showed the best

6    thresholds with 8 cycles compared to all age groups. These differences resulted in

7    better temporal integration for AM detection in these two groups compared to the 7-8

8    and 10-11 years.

9    The fact that the youngest children were not affected by the absolute duration

10   of the sounds suggested that their poorer thresholds did not relate only to a lack of

11   sustained attention but may relate to inefficiencies, e.g., higher internal noise, in

12   processing the temporal-envelope information when few cycles were available.

13   Moreover, the reduced temporal integration observed between 7 and 11 years pointed

14   towards additional constraints imposed by the limited memory buffer involved in AM

15   processing with longer stimulus durations. Computational modelling confirmed that

16   changes in late processing stages over age (e.g., a reduction of internal noise coupled

17   with a reduction in memory noise) better explained changes in temporal integration

18   between 5 and 11 years than changes in the early stages of AM processing. The

19   model, however, did not account correctly for the data at 32 Hz, probably because of

20   the short duration of the sounds and the dominance of the transient response of the

21   simulated AM filters in this condition. The present study overall suggested that

22   processing efficiency for AM continues to develop late into childhood.

23

1    **APPENDIX**

2    The model structure is similar to that used by Wallaert *et al.* (2017). The model

3    had the following stages in sequential order:

4    1. a bank of five gamma-tone filters, one centered at the carrier frequency of the

5    stimulus, and the remaining four centered at 1 and 2 ERB$_N$ above and below the

6    carrier frequency of the stimulus;

7    2. a 'broken-stick' input-output function for the output of the gamma-tone filter tuned

8    to the carrier frequency of the stimulus; the function is linear up to a knee-point of

9    30 dB SPL and compressive (using a power law with an exponent of 0.3) above;

10   3. half-wave rectification of all five frequency channels;

11   4. high-pass filtering (1$^{st}$ order 3 dB/oct roll-off, default cut-off value: 2.5 Hz) of all

12   frequency channels to simulate the effects of short-term adaptation in the

13   modulation domain (Tchorz and Kollmeier, 1999);

14   5. the signal of each frequency channel was passed to a filter-bank (1$^{st}$ order

15   Butterworth filters) with 10 logarithmically-spaced channels tuned between 2 and

16   120 Hz (Moore et al., 2009), each with a $Q$ factor of 1 (Ewert and Dau, 2000) to

17   decompose the modulations of the processed signals, producing 50 channels;

18   6. the model preserved temporal-envelope phase at the output of modulation filters

19   tuned below 6 Hz. The model discarded the envelope phase for channels above 6

20   Hz by passing only the absolute magnitude of the Hilbert transform of the outputs

21   to the following stage;

22   7. two independent Gaussian noises (an 'additive' and a 'memory' noise) were added

23   consecutively to the output of all 50 channels; the first type of noise ("additive

24   noise") had a constant standard deviation (SD) (Dau et al., 1997a); the second type

25   of noise ("memory noise") was additive like the first one, but had an SD which was

1     multiplied by an exponential decay function to model echoic-memory limitation; the

2     addition of this "memory noise" resulted in a weaker representation of the earlier

3     part of the signal than the later and reduced temporal integration of envelope cues

4     (Ardoint et al., 2008; Wallaert et al., 2017), disrupting the representation of longer

5     duration stimuli more than shorter duration stimuli. The decay time constant was

6     fixed at 1.4 s;

7     8. the final decision stage was based on a template matching process (Dau et al.,

8     1997a). The model generated an "internal template" at the start of each staircase

9     with the modulation depth set at the starting value and without any internal noise.

10    The internal template was calculated as the difference between the internal

11    representations of the target and reference stimuli, channel by channel. On each

12    trial, the target and reference stimulus intervals were cross-correlated (channel by

13    channel) with the template. The lags used in the cross-correlation were restricted

14    to ± 1 target modulation cycle. The interval with the largest cross-correlation

15    coefficient (summed across channels) was selected by the model.

16

17    Author Agreement Statement
18
19
20
21
22
23

24    We the undersigned declare that this manuscript is original, has not been
25    published before and is not currently being considered for publication
26    elsewhere.
27
28
29    We confirm that the manuscript has been read and approved by all named
30    authors and that there are no other persons who satisfied the criteria for
31    authorship but are not listed. We further confirm that the order of authors
32    listed in the manuscript has been approved by all of us.

1
2
We understand that the Corresponding Author is the sole contact for the
Editorial process.

He/she is responsible for communicating with the other authors about
progress, submissions of revisions and final approval of proofs

Signed by all authors as follows:

Laurianne Cabrera, Irene Lorenzini, Stuart Rosen, Léo Varnet, Christian Lorenzi

## REFERENCES

Ardoint, M., Lorenzi, C., Pressnitzer, D., Gorea, A., 2008. Investigation of perceptual
constancy in the temporal-envelope domain. J. Acoust. Soc. Am. 123, 1591–
1601. https://doi.org/10.1121/1.2836782

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

Baker, R.J., Rosen, S., 2001. Evaluation of maximum-likelihood threshold estimation
with tone-in-noise masking. Br. J. Audiol. 35, 43–52.

Bartha-Doering, L., Deuster, D., Giordano, V., am Zehnhoff-Dinnesen, A., Dobel, C.,
2015. A systematic review of the mismatch negativity as an index for auditory
sensory memory: From basic research to clinical and developmental
perspectives. Psychophysiology 52, 1115–1130.

Biberger, T., Ewert, S.D., 2016. Envelope and intensity based prediction of
psychoacoustic masking and speech intelligibility. J. Acoust. Soc. Am. 140,
1023–1038.

Buss, E., Lorenzi, C., Cabrera, L., Leibold, L.J., Grose, J.H., 2019. Amplitude modulation detection and modulation masking in school-age children and adults. J. Acoust. Soc. Am. 145, 2565–2575.

Cabrera, L., Varnet, L., Buss, E., Rosen, S., Lorenzi, C., 2019. Development of temporal auditory processing in childhood: Changes in efficiency rather than temporal-modulation selectivity. J. Acoust. Soc. Am. 146. https://doi.org/https://doi.org/10.1121/1.5128324

Cowan, N., 1984. On short and long auditory stores. Psychol. Bull. 96, 341.

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

Dobbins, I.G., Jaeger, A., Studer, B., Simons, J.S., 2012. Use of explicit memory cues following parietal lobe lesions. Neuropsychologia 50, 2992–3003.

Drullman, R., Festen, J.M., Plomp, R., 1994a. Effect of temporal envelope smearing on speech reception. J. Acoust. Soc. Am. 95, 1053.

Drullman, R., Festen, J.M., Plomp, R., 1994b. Effect of reducing slow temporal modulations on speech reception. J. Acoust. Soc. Am. 95, 2670.

Edwards, D.R., Lee, J., Andrews, J., Wong, A., 2008. Contribution of onset/offset information of modulation to amplitude modulation depth discrimination. J. Acoust. Soc. Am. 123, EL111–EL115.

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

Ferguson, A.N., Bowey, J.A., 2005. Global processing speed as a mediator of developmental changes in children's auditory memory span. J. Exp. Child Psychol. 91, 89–112.

Goswami, U., Thomson, J., Richardson, U., Stainthorp, R., Hughes, D., Rosen, S., Scott, S.K., 2002. Amplitude envelope onsets and developmental dyslexia: A new hypothesis. Proc. Natl. Acad. Sci. 99, 10911–10916.

Hall, J.W., 3rd, 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.

Hanks, T.D., Kopec, C.D., Brunton, B.W., Duan, C.A., Erlich, J.C., Brody, C.D., 2015. Distinct relationships of parietal and prefrontal cortices to evidence accumulation. Nature 520, 220–223.

Hoaglin, D.C., Iglewicz, B., 1987. Fine-tuning some resistant rules for outlier labeling. J. Am. Stat. Assoc. 82, 1147–1149.

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

Keller, T.A., Cowan, N., 1994. Developmental increase in the duration of memory for tone pitch. Dev. Psychol. 30, 855.

King, A., Varnet, L., Lorenzi, C., 2019. Accounting for masking of frequency modulation by amplitude modulation with the modulation filter-bank concept. J. Acoust. Soc. Am. 145, 2277–2293.

Lee, J., Bacon, S.P., 1997. Amplitude modulation depth discrimination of a sinusoidal carrier: Effect of stimulus duration. J. Acoust. Soc. Am. 101, 3688–3693.

Levitt, H., 1971. Transformed up-down methods in psychoacoustics. J. Acoust. Soc. Am. 49, 467–477.

Liégeois-Chauvel, C., Lorenzi, C., Trébuchon, A., Régis, J., Chauvel, P., 2004. Temporal envelope processing in the human left and right auditory cortices. Cereb. Cortex 14, 731–740.

Lorenzi, C., Dumont, A., Fullgrabe, C., 2000. Use of temporal envelope cues by children with developmental dyslexia. J. Speech Lang. Hear. Res. 43, 1367.

Massaro, D.W., 1972. Preperceptual images, processing time, and perceptual units in auditory perception. Psychol. Rev. 79, 124.

Moore, B.C., Füllgrabe, C., Sek, A., 2009. Estimation of the center frequency of the highest modulation filter. J. Acoust. Soc. Am. 125, 1075–1081.

Peter, V., Wong, K., Narne, V.K., Sharma, M., Purdy, S.C., McMahon, C., 2014. Assessing spectral and temporal processing in children and adults using temporal modulation transfer function (TMTF), iterated ripple noise (IRN) perception, and spectral ripple discrimination (SRD). J. Am. Acad. Audiol. 25, 210–218.

Rosen, S., 1992. Temporal information in speech: acoustic, auditory and linguistic aspects. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 336, 367–373. https://doi.org/10.1098/rstb.1992.0070

Scott, B.H., Mishkin, M., Yin, P., 2014. Neural correlates of auditory short-term memory in rostral superior temporal cortex. Curr. Biol. 24, 2767–2775.

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

Tchorz, J., Kollmeier, B., 1999. A model of auditory perception as front end for automatic speech recognition. J. Acoust. Soc. Am. 106, 2040–2050.

Varnet, L., Ortiz-Barajas, M.C., Erra, R.G., Gervain, J., Lorenzi, C., 2017. A cross-linguistic study of speech modulation spectra. J. Acoust. Soc. Am. 142, 1976–1989.

Viemeister, N.F., 1979. Temporal modulation transfer functions based upon modulation thresholds. J. Acoust. Soc. Am. 66, 1364.

Viemeister, N.F., Wakefield, G.H., 1991. Temporal integration and multiple looks. J. Acoust. Soc. Am. 90, 858–865.

Walker, B.A., Gerhards, C.M., Werner, L.A., Horn, D.L., 2019. Amplitude modulation detection and temporal modulation cutoff frequency in normal hearing infants. J. Acoust. Soc. Am. 145, 3667–3674.

Wallaert, N., Moore, B.C., Ewert, S.D., Lorenzi, C., 2017. Sensorineural hearing loss enhances auditory sensitivity and temporal integration for amplitude modulation. J. Acoust. Soc. Am. 141, 971–980.

Wallaert, N., Moore, B.C., Lorenzi, C., 2016. Comparing the effects of age on amplitude modulation and frequency modulation detection. J. Acoust. Soc. Am. 139, 3088–3096.

Wallaert, N., Varnet, L., Moore, B.C., Lorenzi, C., 2018. Sensorineural hearing loss impairs sensitivity but spares temporal integration for detection of frequency modulation. J. Acoust. Soc. Am. 144, 720–733.

Witton, C., Stein, J.F., Stoodley, C.J., Rosner, B.S., Talcott, J.B., 2002. Separate influences of acoustic AM and FM sensitivity on the phonological decoding skills of impaired and normal readers. J. Cogn. Neurosci. 14, 866–874.

Yao, J.D., Gimoto, J., Constantinople, C.M., Sanes, D.H., 2020. Parietal cortex is required for the integration of acoustic evidence. Curr. Biol. 30, 3293–3303.