



HAL
open science

Reduced processing efficiency impacts auditory detection of amplitude modulation in children: evidence from an experimental and modeling study

Irene Lorenzini, Christian Lorenzi, Léo Varnet, Laurianne Cabrera

► To cite this version:

Irene Lorenzini, Christian Lorenzi, Léo Varnet, Laurianne Cabrera. Reduced processing efficiency impacts auditory detection of amplitude modulation in children: evidence from an experimental and modeling study. *Hearing Research*, 2024, pp.108982. 10.1016/j.heares.2024.108982. hal-04497338

HAL Id: hal-04497338

<https://hal.science/hal-04497338v1>

Submitted on 13 Dec 2024

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.



Distributed under a Creative Commons Attribution 4.0 International License



Reduced processing efficiency impacts auditory detection of amplitude modulation in children: Evidence from an experimental and modeling study

Irene Lorenzini^{a,c,*}, Christian Lorenzi^b, Léo Varnet^b, Laurianne Cabrera^a

^a Integrative Neuroscience and Cognition Center, CNRS and Université Paris Cité, 45 rue des Saints-Pères, 75006 Paris, France

^b Laboratoire des systèmes perceptifs, Département d'études cognitives, École normale supérieure, Université Paris Sciences et Lettres (PSL University), 29 rue d'Ulm, 75005 Paris, France

^c Laboratoire Ethologie, Cognition, Développement, Université Paris Nanterre, 200 av. de la République, 9200, Nanterre, France

ARTICLE INFO

Keywords:

Amplitude modulation detection
Human auditory development
Processing efficiency
Double-pass consistency task
Modulation filterbank model
Internal noise

ABSTRACT

Auditory detection of the Amplitude Modulation (AM) of sounds, crucial for speech perception, improves until 10 years of age. This protracted development may not only be explained by sensory maturation, but also by improvements in processing efficiency: the ability to make *efficient* use of available sensory information. This hypothesis was tested behaviorally on 86 6-to-9-year-olds and 15 adults using AM-detection tasks assessing absolute sensitivity, masking, and response consistency in the AM domain. Absolute sensitivity was estimated by the detection thresholds of a sinusoidal AM applied to a pure-tone carrier; AM masking was estimated as the elevation of AM-detection thresholds produced when replacing the pure-tone carrier by a narrowband noise; response consistency was estimated using a double-pass paradigm where the same set of stimuli was presented twice. Results showed that AM sensitivity improved from childhood to adulthood, but did not change between 6 and 9 years. AM masking did not change with age, suggesting that the selectivity of perceptual AM filters was adult-like by 6 years. However, response consistency increased developmentally, supporting the hypothesis of reduced processing efficiency in early childhood. At the group level, double-pass data of children and adults were well simulated by a model of the human auditory system assuming a higher level of internal noise for children. At the individual level, for both children and adults, double-pass data were better simulated when assuming a sub-optimal decision strategy in addition to differences in internal noise. In conclusion, processing efficiency for AM detection is reduced in childhood. Moreover, worse AM detection was linked to both systematic and stochastic inefficiencies, in both children and adults.

1. Introduction

The detection of variations in the amplitude of sounds (amplitude modulation, AM) is fundamental for robust auditory processing and is particularly important for speech perception. Indeed, comprehension of speech submitted to severe reduction of spectral cues is maintained as long as AM information is preserved (Remez et al., 1981; Shannon et al., 1995; Van Tasell et al., 1987) and, conversely, degradation of AM information impedes speech comprehension (Drullman, 1995; Dubbelboer and Houtgast, 2007; Houtgast and Steeneken, 1985).

The ability to perceive AM cues and the reliance on them in performing linguistic tasks have been observed since birth. Specifically, newborns' brains have been shown to detect changes in speech sounds (plosive consonants) when the signal is degraded so as to preserve only

the slowest AM cues (< 8 Hz) in each frequency band (Cabrera and Gervain, 2020). These results were shown using electroencephalography, with an oddball design measuring mismatch responses to a change from a frequent to an infrequent consonant stimulus in which the original speech stimuli were degraded to contain only < 8 Hz modulations in amplitude.

Coherently with newborn studies, efficient exploitation of AM cues has been observed in older infants completing behavioral consonant discrimination tasks (Bertoncini et al., 2011; Cabrera et al., 2013), even though some investigations have highlighted greater dependence on faster AM cues, for this kind of task, in 3-month-olds and 6-month-olds as compared with adults (Cabrera et al., 2015; Cabrera and Werner, 2017).

Using non-speech stimuli (amplitude-modulated noises or pure

* Corresponding author at: Laboratoire Ethologie, Cognition, Développement, Université Paris Nanterre, 200 avenue de la République, 92100, Nanterre, France.
E-mail address: ilorenzi@parisnanterre.fr (I. Lorenzini).

tones), Walker et al. (2019) revealed that 3-month-old infants exhibit sensitivity to AM as a function of modulation rate, similarly to adult listeners (i.e., with a low-pass shape when using broadband carriers). Such result strongly suggests that auditory temporal acuity, the limit of the auditory system's ability to follow AM fluctuations as they become faster, is mature early on. However, this study also indicated that infants show poorer AM sensitivity (i.e., higher AM detection thresholds) than adult listeners. Furthermore, investigations conducted with older children have confirmed that auditory sensitivity to AM continues to develop up to 11 years of age (Banai et al., 2011; Buss et al., 2019; Cabrera et al., 2019; Hall and Grose, 1994; Peter et al., 2014).

Two hypotheses have been put forward to explain this slow developmental trajectory of AM detection thresholds: (i) first, it might stem from the maturation of the sensory mechanisms underlying AM extraction: a bank of modulation filters which are centrally implemented and organized in neural sites selectively tuned to specific AM rates (Bacon and Grantham, 1989; Biberger and Ewert, 2016; Dau et al., 1997a; Giraud et al., 2000; Houtgast, 1989; Liégeois-Chauvel et al., 2004); (ii) second, this late development might also depend on immaturity in 'processing efficiency': the ability to make efficient use of the sensory information that has been extracted by the modulation filters when this is available for further central processing. In support of this second hypothesis, computational models have shown that AM perception can be limited by factors such as short-term memory capacities, suboptimal decision making, and internal noise (e.g. Cabrera et al., 2019; 2022). Internal noise, here, is a general concept referring to all forms of neural variability altering both accuracy and consistency. Behaviorally, this is manifested by the fact that, in contrast with ideal observers, human listeners do not provide the same response when presented with the same stimulus multiple times (Green, 1964).

In contradiction with the first hypothesis, that late AM development is linked to late sensory development, previous investigations with both infants and children have shown early maturation of the sensory-processing mechanisms involved in AM processing. Such results have been found when measuring AM detection thresholds at rates from 4 to 128 Hz, using broadband or narrowband noise carriers (Cabrera et al., 2019; Hall and Grose 1994; Walker et al., 2019). To the best of our knowledge, only one study has reported temporal acuity to be worse in children than adults (Buss et al., 2019). In this study, greater differences in AM detection between adults and children (5-to-11-year-olds) were found when AM detection thresholds were measured at higher (e.g. 256 Hz) as opposed to lower rates (e.g. 16, 64 Hz). To further assess the sensory maturity of AM processing, AM maskers have also been employed. Such 'maskers' are competing patterns of AM that can be introduced either in the same frequency region as the carrier ('on-frequency masking') or in a different frequency region ('off-frequency masking'). An 'AM masking effect' is observed when AM detection thresholds are degraded (i.e., elevated) by the introduction of the AM masker compared to a non-masked stimulus condition. Masking effects are thus meaningful, as they are supposed to reflect the selectivity of the human modulations filters tuned to specific AM rates (Bacon and Grantham, 1989; Dau et al., 1997b; Houtgast, 1989). In line with this, previous studies have shown differences in AM masking effects between young children and adults. Specifically, Buss et al. (2019) have observed higher susceptibility to AM masking in 5-year-olds than in adults for off-frequency conditions but not for on-frequency conditions, interpreted as the sign of ongoing maturation of selective listening. Cabrera et al. (2019), in turn, found increasing susceptibility to AM masking between 5 and 7 years of age when comparing AM detection thresholds for deterministic vs non-deterministic carriers (pure tones vs narrowband noises respectively). These results were explained by higher (worse) AM detection thresholds for the younger participants in the deterministic (pure tones) condition, thus reducing AM masking effect in this age group: at 5–6 years of age, children behave as if the stimulus (the pure tone) contains more noise than it really does. In a subsequent modeling study, the authors found that worse AM detection thresholds

in 5-to-6-year-olds were best simulated by implementing adult-like sensory processing (i.e., adult-like modulation frequency selectivity) but with increased levels of internal noise as compared to older children and adults. Acting at the output of the modulation filters, internal noise, here, represents a form of randomness in the auditory system which affects the efficient use of AM information at more central stages. Cabrera et al. (2022) obtained similar results in a modeling study on temporal integration for AM in childhood.

In sum, both experimental investigations and computational modeling of the development of AM detection support the idea that the extraction of AM information is mature from an early age, but that the ability to make an efficient use of this information is sub-optimal in young listeners. Two additional findings can be cited in support of such hypothesis. Firstly, reduced efficiency for AM processing has also been observed in developmental animal studies. For example, juvenile gerbils display higher variability in auditory cortical neuron activity when detecting AM compared with adults, and lower neural firing rates for AM sounds following training for AM detection (Caras and Sanes, 2019). Secondly, improvement of processing efficiency with age has also been shown in humans in other auditory domains. In this regard, Buss et al. (2009) found shallower psychometric functions in school-aged children compared with adults for intensity discrimination, suggesting higher internal noise; Hill et al. (2004) showed that auditory backward masking thresholds are best simulated in 9-to-10-year-olds when using the same temporal window as adults (simulating similar temporal resolution), but higher internal signal-to-masker ratios at the decision stage (simulating poorer processing efficiency).

Despite the converging evidence outlined above, poor processing efficiency in young listeners must be further investigated. Fundamentally, the underlying explanations of this phenomenon remain to be discovered, and studying the cognitive impact of processing efficiency can help to enrich our understanding of the development of core auditory capacities (Sanes and Wooley, 2011). Critically, to date, modeling studies have supported the hypothesis of internal noise impacting the development of AM detection but there is a considerable lack of support of experimental nature, as previous experimental studies only measured detection thresholds.

To fill this gap, we present an experimental and modeling investigation of the hypothesis that reduced AM perception in children might be due to poorer processing efficiency and, specifically, to higher levels of internal noise. In order to address this issue, we first evaluated AM detection and AM masking in a large cohort of 86 children between 6 and 9 years of age as compared to 15 young adults. Specifically, AM detection thresholds were measured in two conditions using (i) a pure tone as carrier (that is, a deterministic carrier with no competing AM fluctuations) or (ii) a narrowband noise as carrier (a stochastic carrier presenting interfering intrinsic random AM fluctuations, cf. Fig. 1 in Varnet and Lorenzi, 2022). This allows for estimation of susceptibility to AM masking as the difference in AM detection thresholds between the two conditions. Then, we assessed processing efficiency for AM detection, a proxy of internal noise in the AM domain, using the 'double-pass psychophysical paradigm'. Introduced by Green (1964) and recently applied by Attia et al. (2021) for AM and FM detection, this paradigm targets variability in perceptual decisions. Precisely, the double-pass consistency task consists in presenting the same set of stimuli two times (i.e., in two 'passes') to a same participant, who takes part in a signal detection task. The participant's responses are expected to vary to a certain extent from the first to the second presentation, and the amount of variation observed, i.e., intersession consistency (estimated as 'Percentage of agreement' in our experimental procedure, see the section 'Procedure' of Task 2) that is the proportion of consistent responses between the two passes on a trial to trial basis, is assumed to relate to the level of internal noise constraining the perceptual decision (Sanes and Wooley, 2011).

Note that we only used noise carriers in this task, as they allow to study the relation between external and internal noise on perceptual

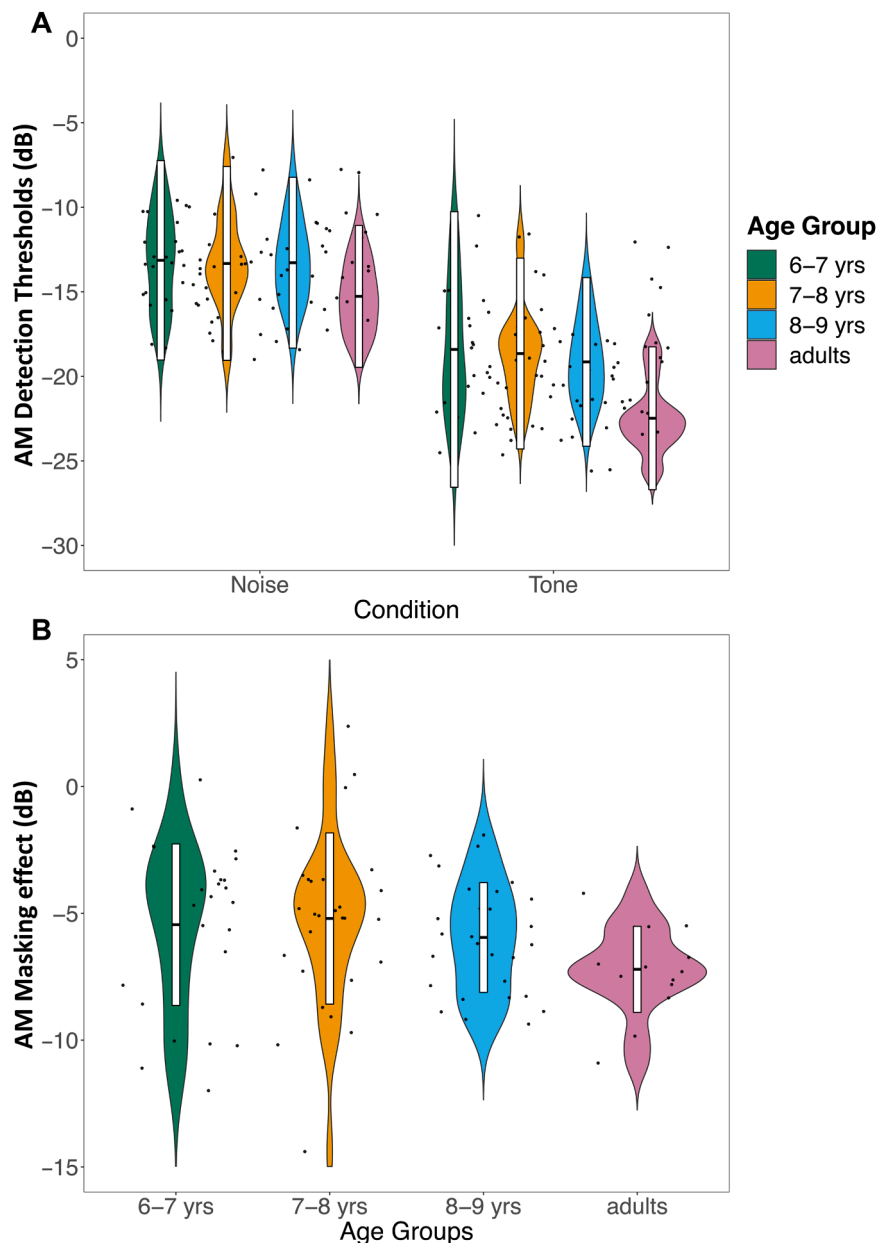


Fig. 1. A. AM-detection thresholds (in dB) for each age group and AM carrier condition (Noise, Tone). Lower thresholds indicate greater AM sensitivity. The boxes in the boxplots represent the 25th, 50th and 75th percentile for each group/condition. The violin plots illustrate the kernel probability density (i.e. the width of the violin area represents the proportion of the data located there). **Fig. 1B.** AM Masking effect (average difference in dB in AMDTs between Tone and Noise conditions) for each age group. The more negative the AM masking, the greater the susceptibility to AM masking.

performance, as listeners tend to show more consistent responses across multiple passes when internal noise is lower than external noise and, conversely, less consistent responses across multiple passes when internal noise is higher than external noise (Attia et al., 2021; Green, 1964).

Based on previous evidence (Cabrera et al., 2019; 2022), we expected to observe in children as compared with adults overall poorer AM sensitivity, but comparable susceptibility to AM masking, and poorer response consistency in the double-pass task.

We then simulated the experimental data using a computational model of the human auditory system based on the modulation filterbank concept, in order to explore whether making vary the variance of internal noise implemented at the output of perceptual AM filters could explain the double-pass results both at the group level and at the individual level. In Cabrera et al. (2019), using a similar modeling

procedure, the variance of internal noise was the only model parameter requiring modification to account for differences in AM detection and masking between young (mean age 5 years) and older listeners (6 to 11 years of age). Here, we expected analogous tendencies for the simulation of the double-pass results at the group level.

2. Experiment 1: AM detection, AM masking and response consistency in middle childhood

2.1. Participants

We recruited 91 children aged between 6 and 9 years at four different schools in Paris and 15 young adults aged between 18 and 27 years. Age groups in our sample, as described in the following sections, are based on French school entry. This means that children born the same calendar

year will be in the same school entry, i.e., Year 1 of primary school will comprise children aged from 6:0 to 6:11 years ('6-to-7-year-olds'); Year 2, 7:0 to 7:11 ('7-to-8-year-olds'); Year 3, 8:0 to 8:11 (8-to-9-year-olds).

According to parents and teachers, the children were developing normally, displayed typical language development and had no special education needs. Opt-out informed consent forms were distributed to the parents, and consent forms were signed by the adult participants as approved by the University ethics committee. Adults filled out a questionnaire detailing their hearing and language history and received a monetary compensation for their time. Receptive vocabulary and non-verbal reasoning were collected as control measures of typical development (*Échelle de Vocabulaire en Images Peabody - Dunn et al., 1993*; the block subtest from the *Wechsler Intelligence Scale for Children, Fourth Edition - Wechsler, 2003*). For both children and adults, absolute auditory sensitivity (i.e., audiometric pure tone thresholds) was obtained with sinusoidal sounds at 250, 500, 1000, 2000, 4000, and 8000 Hz prior to testing. Only data from participants with hearing thresholds \leq 20 dB HL (hearing level) at those frequencies were included.

One child (aged 8 years) was excluded due to experimental error. An additional four children were tested but not included in the final sample, among which: one child (aged 7–8 years) who was excluded due to floor levels of performance and suspicion of positive familial history for speech, language and hearing disorders; one child who refused to cooperate (aged 8-to-9 years) and two children who did not understand the task (aged 6-to-7 years and 8-to-9 years).

The final sample included: 28 6-to-7-year-olds (18 females; mean age = 6.6 years, SD = 0.3), 29 7-to-8-year-olds (17 females; mean age = 7.6 years, SD = 0.3), 29 8-to-9-year-olds (17 females; mean age = 8.5 years, SD = 0.4) and 15 adults (14 females; mean age = 22.5 years, SD = 1.9). The proportion of native bilinguals was negligible: 0/28 6–7-year-olds, 4/29 7–8-year-olds, and 1/29 8–9-year-olds. Moreover, all bilinguals were predominantly hearing and speaking French in daily life. The adults were monolingual French speakers having only learned English at school as a second language.

2.2. Tasks, stimuli and procedure

The participants completed two behavioral tasks articulated in several subtasks. Task 1 measured AM detection thresholds (unmasked AMDTs) and AM masking (masked AMDTs). Task 2 assessed intersession (response) consistency in AM detection through a double-pass procedure using masked AM stimuli.

2.3. Task 1: AM detection and AM masking

2.3.1. Stimuli

All stimuli were generated digitally at a sampling rate of 44.1 kHz using the MATLAB environment (version 2018b). Two carrier conditions were designed to assess AM detection for an 8-Hz target sinusoidal using: (i) a 500-Hz sine tone carrier (Tone condition) and (ii) a 4-Hz wide narrowband noise carrier centered at 500 Hz (Noise condition). We used a total of 500 different noise carriers. All carrier stimuli were 500-ms long, including a 14-ms raised-cosine onset/offset ramps and were generated with a fixed starting phase. Only the target sounds were modulated in amplitude by the 8-Hz sinusoidal modulator: the standard sounds were either unmodulated 500-Hz sine tones or narrowband noises centered at 500 Hz. The expression describing the modulated target was:

$$T(t) = [1 + m \cdot \sin(2\pi \cdot f_m \cdot t + 3\pi / 2)] \cdot c(t)$$

where t is time, m is the modulation depth ($0 \leq m \leq 1$), f_m is the modulation rate and c is the carrier signal.

For the Noise condition, each narrowband noise carrier was generated by adding together five equal-amplitude sine tones with frequencies 498, 499, 500, 501 and 502 Hz. Sounds with “low” values of

envelope Standard Deviation (SD) were selected (SD ranging between 0.059 and 0.069 a.u., arbitrary amplitude units, maximum value of the modulation index spectrum = 0.19, see Fig. 2 in Cabrera et al., 2019). AM spectra of these sound carriers show larger modulation index below about 4 Hz. However, previous studies with children showed AM masking effects using such carriers for AM detection of an 8 Hz modulation target. Using noise carriers with higher inherent envelope fluctuations have been shown to increase AM detection thresholds at 8 Hz modulation close to ceiling for some participants (Cabrera et al., 2019).

All sounds were equated in terms of root-mean-square (RMS) power. The sounds were played diotically through headphones (Sennheiser HD 25-SP II, Sennheiser electronic GmbH and Co. KG, Wedemark, Germany) at 70 dB sound pressure level (SPL), using an external soundcard (DAC audio Audioengine D3). The headphones were calibrated in a double-wall soundproof booth with an artificial ear (4153 6 cc coupler, IEC Standard-60,318–1).

2.3.2. Procedure

Children were tested at school, in a quiet room. As the experiments here reported were part of a larger procedure (also including six speech-in-noise tasks, plus the evaluation of vocabulary skills and non-verbal reasoning), testing lasted 3-weeks in each school. During this period, 7–8-year-olds and 8–9-year-olds completed the whole procedure in five 20-minutes testing sessions, while 5–6-year-olds completed the procedure in six 20-minutes sessions. During each session, a small break was given after 2 or 3 experimental blocks (all lasting 3 to 4 min on average). After each block, the children collected a sticker on a science certificate to keep track of their progress and keep them motivated. Adults were tested in a quiet room at the university and completed all tasks within a single session lasting 1h30, they were allowed to take 10-minute breaks following three or four experimental blocks.

AMDTs were obtained using a two-interval, two-alternative adaptive forced-choice procedure (2I-2AFC) implemented on a touch-screen tablet with a child-friendly interface on Matlab R2018b. On each trial, one interval presented an unmodulated carrier while the target (8-Hz)

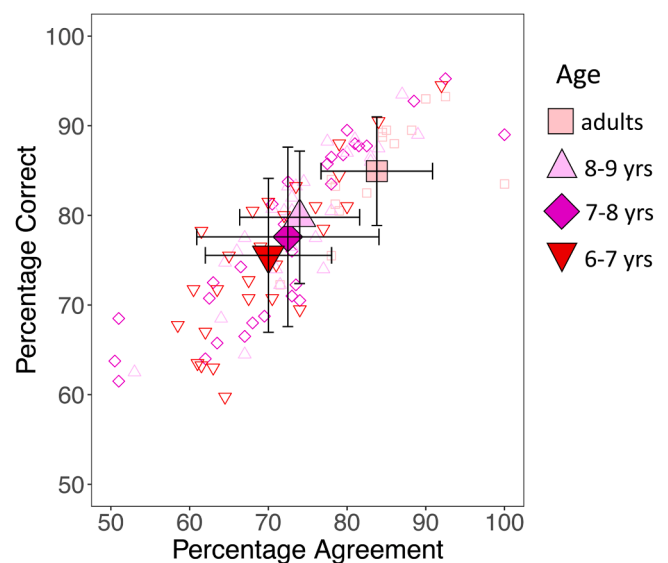


Fig. 2. Percentage Correct and Percentage of Agreement obtained in the double-pass consistency task by each age group (Percentage Correct is averaged on the two passes; Percentage of Agreement is calculated between the two passes). Different symbols and different shades of pink correspond to the PC/PA obtained by the groups (adults = square; 8-to-9-year-olds = point-up triangle; 7-to-8-year-olds = diamond; 6-to-7-year-olds = point-down triangle). Large filled symbols correspond to the average PC/PA of each age group (with bars representing the SD), while individual scores are represented in small open symbols, divided by age group.

sinusoidally amplitude-modulated carrier was presented in the other interval randomly.

For both children and adults, an experimenter sat next to the participant and explained that two animal characters would appear on the screen and produce a sound one after the other. The participants were asked to touch with their finger the character that made “the most buzzy-like sound” (i.e. modulated sound). The interstimulus interval was 500 ms; on each trial, participants had unlimited time to respond and received a feedback (either a smiling or frowning face). The next trial started automatically after the participant’s response.

A transformed and weighted 1-up-1-down adaptive procedure was implemented (see Garcia-Pérez, 1998) and the response on the first trial was always ignored. Step size was changed by a step up of 4 dB after a wrong response and by a step down of 2 dB after a correct response. In the Tone condition, the modulation depth, m , started at -12 dB while it started at -9 dB for the Noise condition. This procedure targeted 76% response correct corresponding to a $d' = 1$ for a 2I-2AFC task.

Prior to testing, both children and adults completed 2 AM detection training runs (one for each condition) to verify whether they were able to perform the task. A participant was excluded if unable to perform the adaptive task (e.g. if, despite repeated attempts and receiving feedback, the participant kept on giving wrong answers or told that she/he could not detect any difference in the stimuli). The scores obtained with these training runs were not used to compute AM thresholds. Then, all participants completed 2 AM detection runs for each condition (thus, each participant completed a total of 3 AM runs under each condition in this task: one training run, followed by two test runs). Each run finished after 50 trials or 8 reversals. Threshold for each run was calculated as the geometric mean of the last 6 reversals. If a difference of more than 3 dB was observed between thresholds obtained in the two experimental runs, or if the psychophysical staircase was not converging, another run was collected and the AMDT in each condition (Tone; Noise) was calculated as the average threshold of the two selected runs (i.e., two runs were selected when both were converging and had no more than 3 dB difference).

AM masking (in dB) was computed as the difference in thresholds between the Tone and the Noise condition.

2.3.3. Task 2: response consistency for AM detection

2.3.3.1. Stimuli. Task 2 measured intersession consistency in the participants using a double-pass procedure. The stimuli were analogous to those used in the Noise condition of Task 1 (narrowband noise carriers centered at 500 Hz) but the modulation depth of the target sounds (m) was individual-dependent, i.e., for each participant the stimuli were played at the individual threshold as estimated in Task 1. Each sound interval presented one noise carrier from a set of 260 different stimuli.

2.3.3.2. Procedure. Using a constant-stimuli procedure, the participants completed a 2I-2AFC where the stimuli were repeated in two passes. One noise carrier was only presented once in each pass. Between the two passes, the trials were exactly the same, that is, the same modulated targets and unmodulated stimulus carriers were presented in the same order within a trial. The order of trial presentation was the same in each pass. Note that intersession consistency in double-pass procedures is independent of whether trial order is always the same or is varied (Hasan et al., 2012).

Each participant completed 2 AM detection passes of 260 trials each, as divided into 10 blocks of 26 trials. At the beginning of each block, 2 “training” trials were played 6 dB above individual threshold (followed by a feedback), then 20 test trials were played at individual threshold, randomly interlaced with 2 catch trials played 4 dB above threshold and 2 additional catch trials played 4 dB below individual thresholds. This was done to capture attention during each block by presenting easier-than-threshold and more difficult trials and these catch trials were not

included into the analyses. No feedback was provided during these 24 trials. Once the 10 blocks of pass 1 were completed, the exact same blocks were repeated in pass 2. The apparatus and interface used were the same as in Task 1. For each participant, the stimuli were played at the individual AM detection threshold as estimated in Task 1 and the level of targeted performance was 76 % ($d' = 1$).

The outcome measures in this task were: Percent Correct (PC) of AM detection in each pass (calculated on the 200 test trials only) and across the two passes (average PC), which signals accuracy in the perceptual decision, and Percent of Agreement (PA) between the response given for the same stimuli in each pass. PA signals consistency in the double-pass and was calculated on a trial-to-trial basis between pass 1 and 2. If different responses were provided for a specific trial between the two passes, a score of 0 was attributed, if the same responses were provided, a score of 1 was attributed. The PA score given to a participant corresponds to these consistency scores averaged across all trials and expressed in%.

3. Results

3.1. Task 1: AM detection and AM masking

We applied the outlier labelling rule to identify potential outliers in all measures (AM detection thresholds and AM masking) and conditions (Tone and Noise condition), i.e., any data value was considered to be an outlier if it lied outside of the following ranges: 3rd quartile + 1.5*interquartile range. 1st quartile – 1.5*interquartile range. No outliers were identified and data had a normal distribution (Kolmogorov-Smirnov tests, all $p > 0.05$).

Data were analyzed using linear mixed models fitted by maximum likelihood (R 3.6.1) using the lme4 package (Bates et al., 2014).¹ In all models, the random effects corresponded to individual participants. For each model, homoscedasticity, linearity and normality were visually inspected on residual plots, and no obvious deviation was detected; p -values were obtained by means of type III analysis of variance (ANOVA) on the models. Tables from I to IV of Appendix A report full statistical details of the models targeting AM detection thresholds and AM masking.

Fig. 1A shows AMDTs (in dB) by age group for the 2 AM carrier conditions. We observe lower (better) AMDTs in the Tone condition compared to the Noise condition at all ages. Moreover, adults show better AMDTs overall than all child groups.

A first statistical model tested the effect of Age Group (6-to-7; 7-to-8; 8-to-9; adults) and Condition (Tone; Noise) entered as fixed effects on AMDTs, entered as the outcome variable. Both Age Group and Condition emerged as significant predictors of AMDTs [Age Group: $F(3101) = 6.024$, $p = 0.0008$; $\eta^2 = 0.152$; Condition: $F(1, 101) = 437.013$, $p < 0.001$, $\eta^2 = 0.812$]. No significant interaction between Age Group and Condition was present [$F(3101) = 1.942$, $p = 0.128$]. Adults obtained better AMDTs than children in both conditions. In all groups, AMDTs obtained with pure tones (Tone condition) were lower (better) than AMDTs obtained with narrowband noises (Noise condition), that included interfering AM fluctuations. Post-hoc Tukey comparisons indicated that AMDTs in adults were significantly lower than in all child groups (all $p < 0.01$). We thus ran a second model only including children’s data, to assess whether AMDTs improved in this restricted age range. The model included AMDTs as the outcome variable, Age (as a linear variable with centered-age values) and Condition (Tone; Noise) as fixed effects. Condition was a significant predictor of AMDTs in children [$F(1,86) = 318.363$, $p < 0.0001$, $\eta^2 = 0.787$], while Age was not [$F(1,86) = 1.150$, $p = 0.287$]. No interaction between Age and Condition was detected [$F(1,86) = 2.471$, $p = 0.120$]. To sum up: AM detection skills were not adult-like by 9 years, but did not significantly improve

¹ The full results of all models are detailed in Appendix A.

with age between 6 and 9 years.

To assess the effect of Age on susceptibility to AM masking, a linear mixed model was run on children and adult data, testing Age Group (6-to-7; 7-to-8; 8-to-9; adults) as a predictor of the Masking effect (the difference between AMDTs in the Tone and Noise condition). Fig. 1B represents susceptibility to AM masking (in dB) by age group. The more negative the AM masking, the greater the susceptibility to AM masking. Results revealed no age-dependent difference in susceptibility to AM masking [$F(3,97) = 1.865$; $p = 0.141$]. Note that, as no interaction between Age and Condition was observed, this implies that Age did not have a significant effect on AM Masking.

Thus, Task 1 showed that both AM detection and susceptibility to AM masking did not significantly improve between 6 and 9 years. However, a significant improvement in AM detection thresholds did emerge from childhood to adulthood and no difference in AM masking was observed. Evidence of early maturity of AM filtering (Walker et al., 2019) encourages the hypothesis, tested in Task 2, that differences in processing efficiency rather than in AM filtering might explain these developmental patterns. This possibility is also supported by previous modeling studies (Cabrera et al., 2019; 2022). Finally, the hypothesis of difference in processing efficiency is coherent with the results of Task 1, where children obtained worse detection thresholds than adults irrespectively of the condition (while better scores in the Tone vs Noise conditions might have suggested worse AM filtering).

3.2. Task 2: Intersession consistency in AM detection

No outliers were identified for Percentage Correct (average on the two passes, PC) nor Percentage Agreement (PA). PA was normally distributed (Kolmogorov-Smirnov test $p > 0.05$). PC distribution differed from normality, having a slightly kurtotic distribution. However, comparison between kurtosis and its SD was considered acceptable for a normal distribution (-1.81 , comprised between $+/-2$, cf. Trochim and Donnelly, 2006). Fig. 2 reports PC/PA scores displayed by each group in the double-pass procedure (PC as averaged on the two passes). Tables from V to VIII of Appendix A report full statistical details of the models targeting PC and PA.

Starting from PC, results were first analyzed in children as compared with adults by using a linear mixed model targeting the fixed effects of Age Group (6-to-7; 7-to-8; 8-to-9; adults), Pass (1 or 2) and Block (10 blocks) on PC. Participants were included as the random effect. The fixed effect of Age Group emerged as significant [$F(3101) = 4.580$; $p = 0.005$; $\eta^2 = 0.119$]. Post-hoc Tukey comparisons revealed that PC scores were significantly lower in 6-to-7- and 7-to-8-year-olds versus adults ($p = 0.004$ and $p = 0.035$ respectively), but not in 8-to-9-year-olds versus adults ($p > 0.05$). This effect can be observed on Fig. 2 (where the SD for PC scores in the 8-to-9-years group is smaller than in younger age groups) and in Figure II of Appendix A, showing PC scores by age groups. The fixed factor of Block was also significant in this model [$F(9, 1919) = 39.962$; $p < 0.0001$; $\eta^2 = 0.157$]. Post-hoc Tukey comparisons explained this effect by, on the one side, a random pattern of differences between blocks (e.g., block 2 yielded better scores than block 7, t -ratio = 3.567, $p = 0.014$; block 8 yielded worse scores than block 9, t -ratio = -3.489 , $p = 0.0178$) and, on the other side, a tendency in all groups to obtain lower PC scores in Block 5 as compared to all other blocks ($p < 0.0001$), which is with all probabilities ascribable to particularly challenging noise carriers randomly selected to constitute this specific block. Importantly, there was no consistent amelioration or decrease of PC paralleling the general progression of blocks from 1 to 10. The fixed effect of Pass was also significant [$F(1,1919) = 4.071$, $p = 0.044$, $\eta^2 = 0.002$], due to improvement from Pass 1 to Pass 2. No significant interaction was detected.

We then ran a linear mixed model only including children data to assess the fixed effects of Age (linear variable with centered-age values), Pass and Block as predictors of PC in children.

Participants were included as the random effect. We observed

significant fixed effects of Age [$F(1,86) = 4.466$; $p = 0.037$; $\eta^2 = 0.049$] and Block [$F(9,1634) = 29.893$; $p < 0.0001$; $\eta^2 = 0.141$]. The effect of Age suggested age-related changes in PC from 6 to 9 years. The effect of Block, as in the model including both children and adults, was due to both random differences and to lower PC scores in Block 5 ($p < 0.0001$ in Post-hoc Tukey comparisons). The fixed effect of Pass was not significant [$F(1,1634) = 3.817$, $p = 0.051$; $\eta^2 = 0.002$], but a significant Age x Pass interaction emerged [$F(1,1634) = 6.282$; $p = 0.012$; $\eta^2 = 0.004$], signaling that older children obtained worse scores in Pass 2 as compared to Pass 1 (interaction slope $\beta = 4.558$, see Figure I in Appendix A). No other significant interaction was detected. Incidentally, this result discourages the hypothesis that high PC scores in older children might result from training or possible learning effects taking place between thresholds estimation and the double-pass procedure. It is important to note that performance (percentage of correct responses) measured with constant stimuli in the double-pass task was expected to match the expected (targeted) performance (76 %) based on the results of the adaptive task. Most participants were found to perform close to this targeted level. However, adults performed generally above the targeted level (85 %) whereas children performed generally below this level (about 70 %). This discrepancy between targeted and actual performance may reflect within-listener variability (that might be higher in children than in adults, and might result from the limited number of reversals, $n = 8$, and runs used in the adaptive task, $n = 2$) or could reflect differences between adults and children in training effects implied by the double-pass task.

PA scores were analyzed, in turn, in a linear mixed model targeting the fixed effects of Age Group (6-to-7; 7-to-8; 8-to-9; adults) and Block (10 blocks) on PA. Participants were included as random effect. In this model, Age Group was a significant predictor of PA [$F(3101) = 7.461$; $p = 0.0001$; $\eta^2 = 0.181$]. As shown by post-hoc Tukey analysis, all child groups obtained significantly lower PA scores than adults (adults versus: 6-to-7-year-olds, $p = 0.0001$; 7-to-8-year-olds, $p = 0.002$; 8-to-9-year-olds, $p = 0.011$). Note, however, that a trend towards improvement in PA from 7-to-8 years of age to 8-to-9 years of age is visible in Fig. 2, paralleling the significant difference observed for PC. The models also revealed a significant fixed effect of Block [$F(9909) = 12.664$, $p < 0.0001$; $\eta^2 = 0.111$], which, as for PC, was related to the lowest PA in block 5 ($p < 0.0001$), showing that block 5 was the block presenting the most difficult stimuli to detect AM from. A significant interaction Age Group x Block was also detected [$F(27,909) = 1.671$, $p = 0.018$; $\eta^2 = 0.047$], however, it was not linked to overall meaningful patterns (but, namely, to 6-to-7- and 7-to-8-year-olds performing better in Block 2 versus 1 and to 8-to-9-year-olds performing better in Block 7 and 10 versus 1).

We then ran a linear mixed model only including children data to assess the fixed effects of Age (linear variable with centered-age values) and Block as predictors of PA. Participants were included as the random effect. Both Age and Block were significant PA predictors [Age: $F(1,86) = 4.780$; $p = 0.031$; $\eta^2 = 0.052$; Block: $F(9774) = 11.446$; $p < 0.0001$; $\eta^2 = 0.117$]. The effect of Block was explained by the same patterns detected by previous models. Additionally, a significant Age x Block interaction emerged, as older children obtained worse PA scores in Block 9 as compared to Block 1 [$F(9774) = 2.024$; $p = 0.034$; interaction slope $\beta = -4.040$]. This effect seems rather incidental, however, as it does not show a meaningful pattern. Importantly, in particular, no consistent improvement was observed alongside block progression (i.e., children did not show strong overall learning effects). Note also that the effect of Block was analogous in children and adults, even though children and adults did not complete double-pass blocks at the same pace (children completing the blocks within several sessions spanning across two weeks, while adults within one session).

Finally, we controlled the relationship between the significant improvements in PC and PA observed between 6 and 9 years of age and AMDTs estimation. This was done to assess whether children with higher (worse) AMDTs have shown improvement in the double-pass

task, and thus, obtain higher PC/PA scores than children with lower (better) AMDTs. To this end, partial correlations controlling for age were ran. The analysis revealed a significant correlation between PC and AM detection in the Noise condition ($r(84) = 0.323$; $p = 0.003$). Table IX of Appendix A reports full statistical details of these analyses. This positive correlation signaled a relationship between worse AMDTs and better PC, revealing a trend where, irrespectively of age, children for whom the double pass was set at smaller AM depth (and thus who showed worse AMDTs) were more accurate in the task. However, crucially, they did not show better coherence in their responses between the two passes as no significant correlation was observed between PA and AMDTs.

In sum, results showed that PC was closer to adult levels by 9 years, while PA was not. Both PC and PA improved with age between 6 and 9 years. Figures IIA and IIB in Appendix A display differences in PC and PA independently in the tested age-groups. Both accuracy and intersession consistency improve with age during childhood. To test whether and to which extent this evolution is related to internal noise, we set up a twofold follow-up modeling study.

4. Models specifications

With a series of simulation studies, we sought to test whether a modulation-filterbank model of the human auditory system could account for age-related improvements of PC and PA. The first simulation study explored whether variations limited to the variance of the internal noise could explain improvements in PC/PA scores at the group level (i. e., all children vs adults, cf. Cabrera et al., 2019). The second simulation study assessed whether the same factor could simulate PC/PA scores at the individual level. In this context, a second parameter – the optimality of the decision mechanism – was also manipulated to improve the model fit.

Simulations were run in the Matlab environment (Dau et al., 1997a, 1997b; Ewert, 2013). PC and PA were simulated using the same stimuli and following the same procedure as human observers. The models used were based on the modulation filterbank concept (Dau et al., 1997a) and they were similar to those used in King et al. (2019), Cabrera et al. (2019, 2022) and Attia et al. (2021). Simulations were articulated into two main processing stages: (i) *front-end processing* (bandpass cochlear filtering, instantaneous, amplitude compression, half-wave rectification, short-term adaptation, AM bandpass filtering aiming to extract temporal-envelope information); and (ii) *a two-component back-end processing stage* including, first, an internal noise component, representing stochastic alterations in the internal representation of AM information and second, a decision strategy module (the decision statistic applied to the (noisy) output of the modulation filters) corresponding to the ‘systematic’ component.

A complete description of the front-end processing stage is provided in Appendix B. Below, we describe the back-end processing stage which is the focus of this study. In accordance with previous studies from our group (Cabrera et al., 2019, 2022), internal noise was modeled as a source of Gaussian noise with a mean of zero directly added to the output of the model. The variance of internal noise was used as a parameter during simulation. Specifically, in each model, the data were simulated by varying the SD of the model internal noise (σ_{in}) to fit the empirical data. The decision strategy module entailed a correlation-based template-matching approach and returned a binary detection decision. A new template computed before each simulation was run, based on a random subset of noisy stimuli. Two types of templates were considered. First, an optimal template was generated corresponding to the most efficient weighting of sensory information. This was achieved by neutralizing both internal and external noise sources during the construction of the template. For this purpose, the SD of internal noise was set to 0. Furthermore, in order to cancel the influence of the intrinsic random fluctuations of the narrowband noise carriers, the optimal template was obtained as the difference between the averaged internal representations of 25 target stimuli (for a suprathreshold

modulation depth of -6 dB SNR) and 25 standard stimuli.² Then, in subsequent simulations, a suboptimal template was also designed. It was generated in the same way as the optimal template, but based on a single pair of target and standard stimuli. Therefore, external noise (the intrinsic random fluctuations of the narrowband noise carrier) had a substantial influence on the resulting template. As a result, the suboptimal template includes features that help discriminate between target and non-target stimuli, as well as features that are specific to the external noise and thus will decrease the performance of the model. Since the template is fixed during the whole experiment, this suboptimality is “systematic” in the sense that it leads to the same responses in the two passes of the experiment. In each trial, the binary decision from the model was obtained by computing the Pearson correlation coefficient between the (noisy) internal representation of each incoming stimuli and the template. The model then selects the interval for which the correlation was highest.

In the first simulation study, double-pass consistency scores (PC and PA) were simulated at the group level, comparing children versus adults (Model A). Children subgroups (all aged between 6 and 9 years) were pooled together here. This choice was motivated by the fact that no difference in AM detection thresholds was found among such subgroups, and allowed us to test replication of previously obtained results (Cabrera et al., 2019). Modulation depth m was set according to each group’s averaged threshold for AM detection in the Noise condition (-15 dB for children, -13 dB for adults) and we varied internal noise SDs so as to best fit the average PC/PA scores obtained by each group (the model was adjusted to the best simultaneous PC/PA fit).

In the second simulation study, we first sought to simulate data at the individual level by varying only the SD of internal noise in the model (Model B). To this end, we simulated PC/PA functions for a range of internal noise SDs and a range of AM depths, m . Internal noise levels were chosen based on the internal noise SD values best fitting group data in the previous simulation; AM depths were chosen so as to mirror the range of AM detection thresholds measured in our participants (-7 to -19 dB). For this second step of modeling, we first ran simulations with a model that, in its final decision stage, used an ‘optimal’ template representation of the target AM, storing the target with perfect accuracy: Model B (cf. Attia et al., 2021). Simulations ran with Model B did not accurately fit all empirical data at the individual level. Therefore, we repeated this simulation study using a model which implemented a ‘suboptimal’ template representation, Model C, that is, a model which stores a representation of the target and non-target stimuli that is not perfectly accurate.

5. Results

5.1. Simulation study 1. effects of varying the internal noise SD on simulated group-level data

Fig. 3 shows the group averaged double pass PC/PA scores as simulated by Model A. The number of trials required to obtain good estimates of PC/PA scores amounted to 500 (10 blocks of 10 trials each). Simulations run without the addition of internal noise to the model were not accurate and, specifically, provided better scores than the real group-level scores. On the contrary, simulating a decrease in internal noise SD as the only variation in the model accurately captured the differences observed at the group level in our experimental data. Specifically, the level of internal noise fitting participant data was larger by a factor of two for children than for adults [500 m.u. (model units) for adults versus 1000 m.u. for children]. Prediction error (mean difference in absolute value between empirical and simulated data) was 1 % for PC and 7 % for PA in adults, 2 % for PC and 8 % for PA in children.

² Previous simulations showed that increasing the number of target and standard stimuli beyond 25 did not improve model performance.

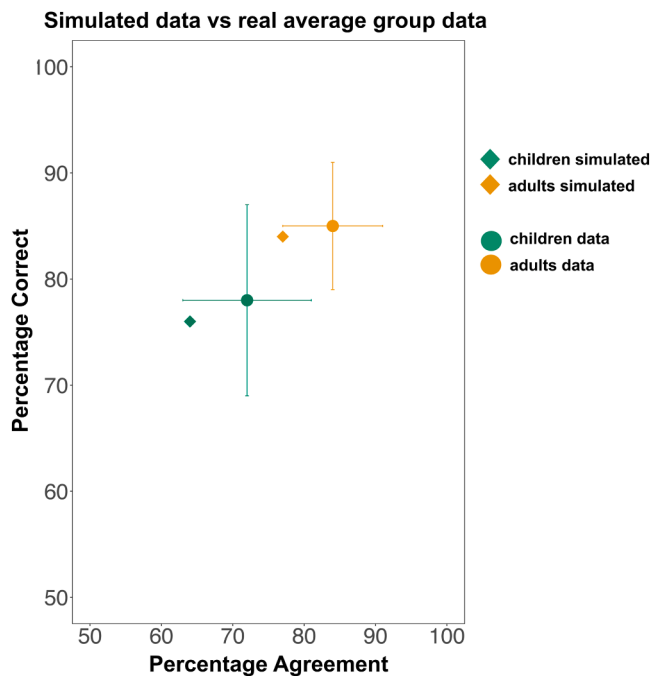


Fig. 3. Real mean PC and mean PA (and SD) for each age group are represented along with the simulated PC/PA scores of Model A. Children real scores are represented by a green dot; children simulated scores by a green diamond. Adult real scores are represented by a yellow dot; adult simulated scores by a yellow diamond.

5.2. Simulation study 2. effects of varying the internal noise SD on the simulated individual data

We then simulated individual PC/PA data points. The SDs of internal noise implemented in the models were chosen starting from the values that best simulated empirical data at the group level (500–1500 m.u.). Additional internal noise levels were then progressively added in 200 m.u. discrete steps to improve model fit. For each internal noise level, PC and PA scores were simulated for 8 AM depths comprised between -20 dB and -6 dB using 2 dB discrete steps, mirroring the range of AM detection thresholds measured in real participants. The number of simulated trials per pass was 500 for each level of internal noise instead of 100 total trials per pass for each real participant (i.e., 10 blocks of 50 trials per 8 AM depths, amounting to 4000 total trials per pass for each level of internal noise).

Fig. 4 shows empirical PC/PA data points (in pink color scale with open symbols) projected together with simulated PC/PA data points (in blue color scale with filled circles). A selection of four levels of internal noise from the stimulation study is presented. The extremes of this four-step consist, on the one side, in extremely low levels of internal noise, which are not psychologically plausible and are shown to contextualize human performance in comparison to the performance of an ideal observer; on the other side, in very high levels of internal noise. The internal noise values were chosen based on the average values of internal noise needed to simulate group performance in the first simulation study (Model A). The reader should appreciate the degree of overlap between the pink open symbols (empirical data) and the blue filled symbols (simulated data); a high degree of overlap indicates a good fit of the simulation to the behavioral data. It appears in Fig. 4 that simulated data are close to the maximum theoretical value of PC for a given PA level only for data points generated with the lower levels of internal noise (either 10 m.u.) and for data points generated with higher levels of internal noise in association with larger (easier) AM depths (between -6 and -10 dB for $\sigma_{\text{int}} = 500$ and 1300 m.u., and limited to -6 dB for $\sigma_{\text{int}} = 2100$ m.u.).

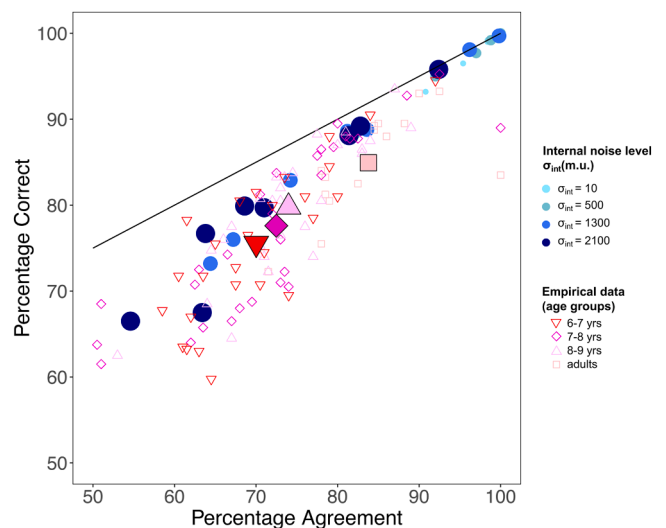


Fig. 4. Individual PC/PA data points for real data (open symbols using pink color scale) and simulated data (filled circles using blue color scale and size according to the internal noise SD) obtained with Model B, using an optimal template and varying levels of internal noise. Average PC/PA scores for each experimental group are represented using filled larger symbols. Each circle represents PC/PA simulated for one out of 8 values of AM depth (ranging from -19 dB to -7 dB in 2 dB discrete steps) and one out of four internal noise levels (dark blue and largest size corresponding to the highest level). The line defined by the theoretical maximum of the PC/PA function is shown ($PC = 100 - (100 - PA)/2$).

On the contrary, real data deviated from this line, due to the fact that some participants had low PC scores that were systematic (and relatively higher PA), i.e. not only caused by the presence of internal noise. In other words, some participants were more consistent than would be expected based on their low PC scores. This point might be explainable by the introduction of a form of systematic suboptimality. We thus set up a third model (Model C) to explore this hypothesis.

5.3. Effects of varying the internal noise SD and template optimality on the simulated individual data

Simulations were repeated with Model C, where internal noise SDs were varied and, in addition, the decision module of the model used a suboptimal template (see Model specifications, pp. 22–23). PC and PA were simulated for the same range of internal noise levels and AM depths as in the previous simulation, yielding a total of 5000 simulated trials.

Despite increasing precision, Model C remained inaccurate for some participants. In other words, age-related decrease in internal noise levels associated with poorer template selectivity did not entirely capture all individual data. Particularly, it left out individuals close to the maximum theoretical line. Fig. 5 shows a selection of seven internal noise levels from the simulation study best illustrating the PC/PA relationship in Model C. As in Fig. 4, degree of overlap between the pink symbols (empirical data) and the blue symbols (simulated data) indicates a good fit of the simulation to the behavioral data. In Fig. 4, this is observed only in the upper right region (close to the maximum theoretical line), while in Fig. 5, this is observed mainly in the lower left region but notably not approaching the maximum theoretical line.

Despite failing to provide accurate simulation of the data, Model C, implementing both variations in internal noise levels and suboptimal template matching, allowed to highlight the theoretical limitations of Model B, where only internal noise levels were varied and template matching was optimal (optimal template). This simulation study showed that, in order to account for the dispersion of individual data points in a double-pass task, a systematic suboptimal component is needed.

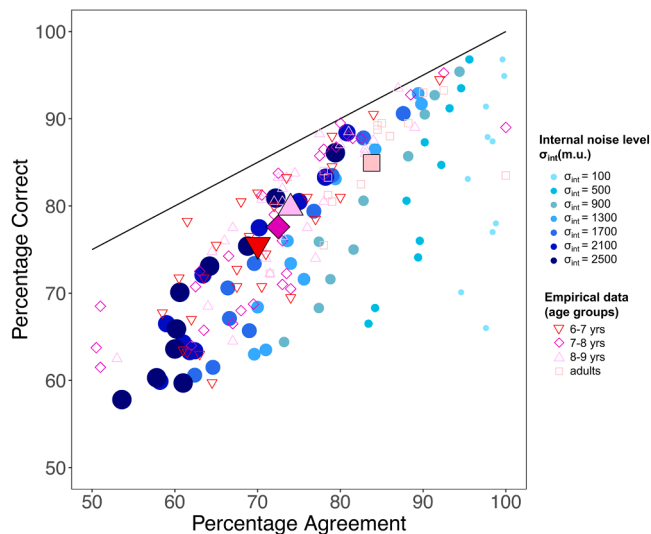


Fig. 5. Individual PC/PA data points for real data (open symbols as a function of age group using pink color scale) and simulated data (filled circles using blue color scale and size according to internal noise level) obtained with Model C, using a suboptimal template and varying levels of internal noise.

To summarize, simulations based on the exclusive manipulation of internal noise SDs did not accurately account for our empirical data at the individual level, nor did simulations where processing efficiency was further degraded by the implementation of a suboptimal template in the model. Importantly, it can be remarked that these two simulation sets produced complementary patterns. Precisely, simulations run with the optimal template (Fig. 4) left out data points where PA scores were higher than PC scores. Conversely, simulations ran with the suboptimal template (Fig. 5) captured the data points that the model failed to capture in the first study, but were less precise in simulating data points falling close to the theoretical maximum of the function.

Thus, the implementation of a suboptimal template proves useful; however, further studies are needed to identify the degree of suboptimality that is necessary to provide a full account of inter-individual variability. These results show that a form of systematic suboptimality at the decision stage is needed, alongside with varying levels of internal noise, to capture the PC/PA space of variation at the individual level.

The observed differences in the relationship between optimal versus suboptimal template and variations in internal noise levels are psychologically plausible. They are discussed below, in light of the developmental trends observed for PC and PA scores.

6. Discussion

6.1. Development of AM detection and AM masking from 6 to 9 years of age and in children versus adults

AM detection was tested using carriers with or without inherent random competing AM fluctuations (i.e., narrowband noises in the Noise condition, versus pure tones in the Tone condition). As expected based on previous studies (Hall and Grose, 1994; Banai et al., 2011; Peter et al., 2014; Buss et al., 2019; Cabrera et al., 2019), AMDTs were not fully adult-like by 9 years of age. We also observed that the thresholds did not significantly improve with age between 6 and 9 years. This is consistent with the findings in Cabrera et al. (2019), where only children below 6 years of age showed worse AMDTs than older children.

Concerning AM masking, our results suggest that the response of AM filters is comparable between children and adults, as masking effects do not differ significantly between the groups. Since the modulation-noise masker used here has a triangular AM spectrum (Lawson & Uhlenbeck, 1950), its contribution to masking depends on the selectivity of the AM

filter centered on the 8 Hz target modulation. Specifically, a wider AM filter will result in a lower envelope signal to-noise ratio and, in turn, in lower (worse) masked AM-detection thresholds. However, the modeling study by Cabrera et al. (2019) showed that widening AM filters by a factor of 2 would only produce a rather small 1–2 dB increase in the masking effect. In the present results, we did not observe significant age differences in the masking effect. Therefore, the present results suggest that - at least - the AM filter width is not widened by more than a factor two for children. This same study (Cabrera et al., 2019), also reported no difference in susceptibility to AM masking between children and adults, and no significant developmental change between 6 and 9 years of age.

In sum, the results of our first experimental study were not compatible with an explanation based only on the development of deterministic sensory factors. This pattern motivated a subsequent experimental exploration of processing efficiency as a factor potentially affecting the development of AM detection in the same group of participants.

6.2. Intersession consistency in AM detection

6.2.1. Experimental investigation of processing efficiency in AM detection

This is - to the best of our knowledge - the first study to experimentally assess AM detection consistency in childhood with a large group of participants ($N = 86$). We tested the hypothesis, supported by previous simulation studies (Cabrera et al., 2019, 2022; Hill et al., 2004), that processing efficiency for AM detection improves during childhood due to a decrease in internal noise. Specifically, participants completed a double-pass task measuring both percent correct (PC) and percent agreement (PA) in AM detection. The results showed significant differences in both PC and PA between children and adults as well as significant improvements in PC and PA between 6 and 9 years of age.

According to Green (1964), lower levels of PA may reflect higher levels of internal noise in participants. Mechanisms such as stimulus expectancies might also affect PA. However, these mechanisms were controlled in the present psychoacoustic procedure: in the double-pass task, the presentation of the target interval was fully random, and different carriers were used on each interval; no feedback was provided after the test trials, and four catch trials (2 played at +2 dB from individual threshold, and 2 played at -2 dB) were randomly played in each block to keep the participants engaged in the task and reducing expectancies. Moreover, age group variations in PA were not linked to participant fatigue, as shown by the effect of Block, which did not reveal significant improvements in PC/PA for the initial as compared to the final blocks of the double-pass.

We thus hypothesized that lower PA scores observed in children as compared to adults ($p = 0.0001$, see Fig. 3) might relate to higher internal noise levels in childhood, and tested this hypothesis in two modeling studies.

Note that PC and PA are by nature strongly related. However, there is no one-to-one relationship between PC and PA (cf. Figs. 4 and 5). In other words, a same PC value can to some extent be linked to different PA values given by different amounts of deterministic or stochastic limitations. This is apparent on Figs. 2, 4 and 5. Formulated in terms of the modulation-filterbank model, the deterministic aspects correspond to limits in the sensory processing of AM, while the stochastic aspects to internal noise. Thus, with our second modeling study we sought to better understand why higher or lower PA values can be linked to a same PC value.

6.2.2. Computational investigation of processing efficiency in AM detection

Changes in processing efficiency between children and adults for AM detection were simulated in a series of modeling studies. Using the modulation filterbank model, we simulated PC/PA scores obtained by our participants in the double-pass consistency task by altering the model's processing efficiency in two different ways: (i) increasing the variance of an additive internal noise at the output of the modulation

filters and (ii) decreasing the optimality of the template used by the decision module.

We first simulated PC/PA functions at the group level (Model A) and obtained accurate simulation of average group-level data (children versus adults) by only varying the level of internal noise of the model. The level of internal noise used to capture group-level PC/PA in adults was two times lower as compared to the children group-level data. It can be noted that the magnitude of internal noise best-fitting participant data in this study does not coincide with the values observed in [Cabrera et al. \(2019\)](#), where the level of internal noise best simulating group-level performance was a factor of 10 smaller for adults compared to the youngest children tested (aged 5–6 years). This discrepancy seems to depend on the differences in the tasks and stimuli used in the two studies, and namely: (i) the use of a 3AFC modulation detection task in a [Cabrera et al. \(2019\)](#) as compared to a double-pass consistency task in the present study; (ii) the use of 3 AM carriers and 3 target AM rates in [Cabrera et al. \(2019\)](#) as compared to narrowband noise carriers centered at 500 Hz in the present study. We conclude that this simulation study demonstrates the usefulness of the modulation filterbank model, and that changes in the model associated to internal noise levels mimic developmental trends in processing efficiency for AM detection.

Subsequently, in our second and third simulation studies, we aimed to capture individual relationships between PC/PA scores. Firstly, we implemented Model B, using an optimal template, and only varied the internal noise levels in the model. Secondly, we repeated the modeling with Model C, which included a sub-optimal template together with variations of internal noise. Simulations modifying only the internal noise parameter (Model B), or simulations both modifying internal noise and implementing a sub-optimal template (Model C) were not accurate enough to entirely capture individual variability, when taken separately. However, when considered in combination, these two models captured the majority of individual data points. This suggests that some of the participants' data were better simulated by elevated internal noise only, while others were better simulated by a combination of elevated internal noise and sub-optimal template building. It seems, thus, that multiple (stochastic and deterministic) factors have to be considered to successfully simulate individual data in an AM detection task. In particular, Model B, which only varied internal noise, yielded highly correlated PC/PA that decreased as the levels of internal noise increased. This model captured accurately individual data points that were close to the theoretical maximum line relating PA and PC scores, indicating that, for these participants, the only limiting factor of performance was internal noise (or, rather, that the effect of the systematic component was negligible compared to the effect of the stochastic component). In Model C (which included variations in internal noise levels and a suboptimal template), lower internal noise levels led to worse PC scores (increasing errors) for lower (i.e., more challenging) AM depths, but not to worse PA scores (see [Fig. 5](#), internal noise = 100 m.u.). Conversely, higher internal noise levels led to a decrease in both PC and PA for lower AM depths (see [Fig. 5](#), internal noise = 2100 m.u.). This combination helped us to better simulate participant falling below the theoretical maximum line. In conclusion, a systematic suboptimal component operating at the decisional level needs to be included in order to accurately capture inter-individual variability in AM detection measured with the double-pass consistency task. The degree of variation needed for this component to fit individual data with full accuracy, as associated to varying levels of internal noise, is left open for future studies.

To explore the impact of age on the observed variability in systematic and non-systematic forms of processing inefficiency, the simulated data obtained with the two models can be compared to the developmental trends observed in PA/PC scores as follows. While obtaining the same PA scores as younger children, older children (aged of 8-to-9-years) made fewer errors through the two passes (as indicated by better PC scores). Therefore, we can speculate that results in this specific age group may relate to the fact that 8–9-year-old children have already built an optimal template that is both stable and accurate (at least more

stable and accurate than in younger children), but that their level of internal noise is still higher than in adults and more similar to younger children.

In sum, a decrease in internal noise level with age for AM detection seems to interact with the development of more accurate and stable internal templates of the target AM to be detected. Future experimental studies should further investigate how children build and represent an AM template and whether these mechanisms change over age, especially between 5 and 10 years, as AM detection seems more limited by processing inefficiencies rather than temporal modulation selectivity in this age range ([Cabrera et al., 2019](#); [Cabrera et al., 2022](#)). This will help to better characterize inter-individual variability and developmental effects in AM detection tasks.

6.3. Limitations of this study

It seems important to underline that follow-up experimental studies might extend our results with reference to different AM rates (even though a large set of AM modulation rates has already been tested in former investigations, [Hall and Grose, 1994](#)). Younger children should also be included in the experimental samples, in order to test consistency in AM detection in a wider developmental perspective. Nevertheless, this seems difficult in reason of the attentional effort required in the double-pass procedure. Less effortful experimental paradigms should be designed to extend the investigation to children under 6 years of age.

As to our modeling studies, a main limitation consists in the fact that the relationship between internal noise levels and the suboptimal template cannot be interpreted directly. The processing inefficiencies that are related to these two factors might be related, that is, errors in the stimulus representation due to high levels of internal noise would lead to errors in the target template formed based on those stimuli. However, the present study does not allow to investigate such relationship. It is important for future experimental studies to explore the development of AM template representation in childhood. Finally, other model parameters might be modified to further improve fit of the real data, such as the addition of a memory noise that has been shown to influence temporal integration in AM processing ([Cabrera et al., 2022](#)).

7. Conclusions

The main goal of the present study was to test whether the development of AM detection capacities between children and adults may relate to differences in AM processing efficiency. AM detection thresholds and susceptibility to AM masking were measured using stochastic (narrowband noises) and deterministic (pure tones) carriers. Children aged from 6 to 9 years showed significantly poorer AM-detection thresholds compared to young adults for an 8 Hz sinusoidally-modulated target, but no significant improvement was observed between 6 and 9 years of age. In line with previous studies, no difference in AM masking was observed between 6-to-9 year-old children and adults. Results obtained in a double-pass task measuring variability in AM detection performance for the same participants tested at threshold, revealed age-related improvement in both detection performance and detection consistency. Follow-up simulations using a model of the human auditory system accounted relatively well for the performance changes in both PC and PA scores between the child and the adult groups by simulating a decrease in internal noise levels with age. However, a second series of models revealed that (systematic) suboptimal decision making needs to be implemented additionally with changes in internal noise levels to obtain accurate AM detection performance simulations at the individual level. Overall, this investigation provided new experimental and modeling evidence to the hypothesis that the development of AM detection (and thus, more broadly, of temporal auditory processing) during childhood is related to reduced processing efficiency. Future studies are needed to systematically assess the mechanisms underlying individual differences in AM detection skills over development.

CRedit authorship contribution statement

Irene Lorenzini: Data curation, Writing – review & editing, Writing – original draft, Investigation, Formal analysis. **Christian Lorenzi:** Writing – review & editing, Conceptualization. **Léo Varnet:** Writing – review & editing, Software, Formal analysis. **Laurianne Cabrera:** Writing – review & editing, Validation, Supervision, Software, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare no financial or non-financial conflict of interest.

Data availability

Data will be made available on request.

Acknowledgments

This work was supported by the ANR-17-CE28-0008 (Agence Nationale de la Recherche) to Laurianne Cabrera and by the MSCA-INCODE (Grant Agreement n. 895484) to Irene Lorenzini. The authors wish to thank Emmanuel Ponsot for valuable insights on the analysis and modelling of the data; all the children that took part into the project and all the teachers that helped realizing the testing sessions; Paula Perrineau and Alicia Lefebvre for help with testing; Sarah Attia and Andrew King for help with setting up the experimental design. The authors declare no conflict of interest. Leo Varnet and Christian Lorenzi were supported by the FrontCog ANR-17-EURE-0017 grant.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.heares.2024.108982](https://doi.org/10.1016/j.heares.2024.108982).

References

- Attia, S., King, A., Varnet, L., Ponsot, E., Lorenzi, C., 2021. Double-pass consistency for amplitude-and frequency-modulation detection in normal-hearing listeners. *J. Acoust. Soc. Am.* 150 (5), 3631–3647.
- Bacon, S.P., Grantham, D.W., 1989. Modulation masking: effects of modulation frequency, depth, and phase. *J. Acoust. Soc. Am.* 85 (6), 2575–2580.
- Banai, K., Sabin, A.T., Wright, B.A., 2011. Separable developmental trajectories for the abilities to detect auditory amplitude and frequency modulation. *Hear. Res.* 280 (1–2), 219–227.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823*.
- Bertoncini, J., Nazzi, T., Cabrera, L., Lorenzi, C., 2011. Six-month-old infants discriminate voicing on the basis of temporal envelope cues (L). *J. Acoust. Soc. Am.* 129 (5), 2761–2764.
- Biberger, T., Ewert, S.D., 2016. Envelope and intensity based prediction of psychoacoustic masking and speech intelligibility. *J. Acoust. Soc. Am.* 140 (2), 1023–1038.
- Buss, E., Hall III, J.W., Grose, J.H., 2009. Psychometric functions for pure tone intensity discrimination: slope differences in school-aged children and adults. *J. Acoust. Soc. Am.* 125 (2), 1050–1058.
- 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 (4), 2565–2575.
- Cabrera, L., Gervain, J., 2020. Speech perception at birth: the brain encodes fast and slow temporal information. *Sci Adv* 6 (30), eaba7830.
- Cabrera, L., Werner, L., 2017. Infants' and adults' use of temporal cues in consonant discrimination. *Ear Hear* 38 (4), 497.
- Cabrera, L., Bertoncini, J., Lorenzi, C., 2013. Perception of speech modulation cues by 6-month-old infants. *J. Speech, Lang., Hear. Res.* 56 (6), 1733–1744.
- Cabrera, L., Lorenzi, C., Bertoncini, J., 2015. Infants discriminate voicing and place of articulation with reduced spectral and temporal modulation cues. *J. Speech, Lang., Hear. Res.* 58 (3), 1033–1042.
- 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 (4), 2415–2429.
- Cabrera, L., Lorenzini, I., Rosen, S., Varnet, L., Lorenzi, C., 2022. Temporal integration for amplitude modulation in childhood: interaction between internal noise and memory. *Hear. Res.* 415, 108403.
- Caras, M.L., Sanes, D.H., 2019. Neural variability limits adolescent skill learning. *J. Neurosci.* 39 (15), 2889–2902.
- 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 (5), 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 (5), 2906–2919.
- Drullman, R., 1995. Speech intelligibility in noise: relative contribution of speech elements above and below the noise level. *J. Acoust. Soc. Am.* 98 (3), 1796–1798.
- Dubbelboer, F., Houtgast, T., 2007. A detailed study on the effects of noise on speech intelligibility. *J. Acoust. Soc. Am.* 122 (5), 2865–2871.
- Dunn, L.M., Thériault-Whalen, C.M., Dunn, L.M., 1993. Peabody Picture Vocabulary test-revised: French adaptation. *Psychon*, Toronto, Canada.
- Ewert, S.D., 2013. AFC-A modular framework for running psychoacoustic experiments and computational perception models. *Proceedings of the international conference on acoustics AIA-DAGA*, 2013, pp. 1326–1329.
- García-Pérez, M.A., 1998. Forced-choice staircases with fixed step sizes: asymptotic and small-sample properties. *Vision Res.* 38 (12), 1861–1881.
- Giraud, A.L., Lorenzi, C., Ashburner, J., Wable, J., Johnsrude, I., Frackowiak, R., Kleinschmidt, A., 2000. Representation of the temporal envelope of sounds in the human brain. *J. Neurophysiol.* 84 (3), 1588–1598.
- Green, D.M., 1964. Consistency of auditory detection judgments. *Psychol. Rev.* 71, 392–407.
- Hall III, J.W., Grose, J.H., 1994. Development of temporal resolution in children as measured by the temporal modulation transfer function. *J. Acoust. Soc. Am.* 96 (1), 150–154.
- Hasan, B.A.S., Joosten, E., Neri, P., 2012. Estimation of internal noise using double passes: does it matter how the second pass is delivered? *Vision Res.* 69, 1–9.
- Hill, P.R., Hartley, D.E., Glasberg, B.R., Moore, B.C., Moore, D.R., 2004. Auditory processing efficiency and temporal resolution in children and adults. *J. Speech, Lang., Hear. Res.* 47, 1022–1029.
- Houtgast, T., Steeneken, H.J., 1985. A review of the MTF concept in room acoustics and its use for estimating speech intelligibility in auditoria. *J. Acoust. Soc. Am.* 77 (3), 1069–1077.
- Houtgast, T., 1989. Frequency selectivity in amplitude-modulation detection. *J. Acoust. Soc. Am.* 85 (4), 1676–1680.
- 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 (4), 2277–2293.
- Lawson, J.L., Uhlenbeck, G.E., 1950. *Threshold signals*. Chicago.
- Liégeois-Chauvel, C., Lorenzi, C., Trébuchon, A., Régis, J., Chauvel, P., 2004. Temporal envelope processing in the human left and right auditory cortices. *Cerebral Cortex* 14 (7), 731–740.
- 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 (02), 210–218.
- Remez, R.E., Rubín, P.E., Pisoni, D.B., Carrell, T.D., 1981. Speech perception without traditional speech cues. *Science* 212 (4497), 947–950.
- Sanes, D.H., Woolley, S.M., 2011. A behavioral framework to guide research on central auditory development and plasticity. *Neuron* 72 (6), 912–929.
- Shannon, R.V., Zeng, F.G., Kamath, V., Wygonski, J., Ekelid, M., 1995. Speech recognition with primarily temporal cues. *Science* 270 (5234), 303–304.
- Trochim, W.M.K., Donnelly, J.P., 2006. *The Research Methods Knowledge Base: Atomic Dog*. Cincinnati, OH.
- Van Tasell, D., Soli, S., Kirby, V., Widin, G., 1987. Speech waveform envelope for consonant recognition. *JASA* 82, 1152–1161.
- Varnet, L., Lorenzi, C., 2022. Probing temporal modulation detection in white noise using intrinsic envelope fluctuations: a reverse-correlation study. *J. Acoust. Soc. Am.* 151 (2), 1353–1366.
- 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 (6), 3667–3674.
- Wechsler, D., 2003. *Wechsler intelligence scale for children—Fourth. Edition (WISC-IV)*.