Vocal Emotion Recognition in School-Age Children With Hearing Aids

Laura Rachman,[1,2,3] Gizem Babaoğlu,[1,4] Başak Özkişi Yazgan,[1,4] Pınar Ertürk,[1,4] Etienne Gaudrain,[5]
Leanne Nagels,[1,3] Stefan Launer,[6] Peter Derleth,[6] Gurjit Singh,[6,7,8] Frédérick Uhlemayr[6]
Monita Chatterjee,[9] Esra Yücel,[4] Gonca Sennaroğlu,[4] and Deniz Başkent[1,3]

Objectives: In individuals with normal hearing, vocal emotion recognition continues to develop over many years during childhood. In children with hearing loss, vocal emotion recognition may be affected by combined effects from loss of audibility due to elevated thresholds, suprathreshold distortions from hearing loss, and the compensatory features of hearing aids. These effects could be acute, affecting the perceived signal quality, or accumulated over time, affecting emotion recognition development. This study investigates if, and to what degree, children with hearing aids have difficulties in perceiving vocal emotions, beyond what would be expected from age-typical levels.

Design: We used a vocal emotion recognition test with non–language-specific pseudospeech audio sentences expressed in three basic emotions: happy, sad, and angry, along with a child-friendly gamified test interface. The test group consisted of 55 school-age children (5.4 to 17.8 years) with bilateral hearing aids, all with sensorineural hearing loss with no further exclusion based on hearing loss degree or configuration. For characterization of complete developmental trajectories, the control group with normal audiometric thresholds consisted of 86 age-matched children (6.0 to 17.1 years), and 68 relatively young adults (19.1 to 35.0 years).

Results: Vocal emotion recognition of the control group with normal-hearing children and adults improved across age and reached a plateau around age 20. Although vocal emotion recognition in children with hearing aids also improved with age, it seemed to lag compared with the control group of children with normal hearing. A group comparison showed a significant difference from around age 8 years. Individual data indicated that a number of hearing-aided children, even with severe degrees of hearing loss, performed at age-expected levels, while some others scored lower than age-expected levels, even at chance levels. The recognition scores of hearing-aided children were not predicted by unaided or aided hearing thresholds, nor by previously measured voice cue discrimination sensitivity, for example, related to mean pitch or vocal tract length perception.

Conclusions: In line with previous literature, even in normal hearing, vocal emotion recognition develops over many years toward adulthood, likely due to interactions with linguistic and cognitive development. Given the long development period, any potential difficulties for vocal emotion recognition in children with hearing loss can only be identified with respect to what would be realistic based on their age. With such a comparison, we were able to show that, as a group, children with hearing aids also develop in vocal emotion recognition, however, seemingly at a slower pace. Individual data indicated a number of the hearing-aided children showed age-expected vocal emotion recognition. Hence, even though hearing aids have been developed and optimized for speech perception, these data indicate that hearing aids can also support age-typical development of vocal emotion recognition. For the children whose recognition scores were lower than age-expected levels, there were no predictive hearing-related factors. This could be potentially reflecting inherent variations related to development of relevant cognitive mechanisms, but a role from cumulative effects from hearing loss is also a possibility. As follow-up research, we plan to investigate if vocal emotion recognition will improve over time for these children.

Key words: Children, Development, Hearing aids, Hearing loss, Vocal emotion.

(Ear & Hearing 2025;XX;00–00)

This article has received OSF badges for Open Data and Open Material.

INTRODUCTION

Communicating emotions is an important aspect of social interactions (Picou et al. 2018). Studies on the development of vocal emotion recognition in children with normal hearing (NH) have shown that cortical specialization of vocal emotion processing starts in infancy (Blasi et al. 2011; Cheng et al. 2012). However, overall vocal emotion recognition abilities continue to develop throughout childhood and adolescence, especially if children need to rely on prosodic affective cues only, for example, when the linguistic affective content is not available (Aguert et al. 2013; Sauter et al. 2013; Chronaki et al. 2015; Grosbras et al. 2018; Nagels et al. 2020b; Amorim et al. 2021; Filippa et al. 2022).

For children with hearing loss and hearing devices, prosodic cues may be altered, and vocal emotion recognition and its development may differ from that of NH. Various acoustic cues can convey vocal emotions, such as mean fundamental frequency (F0), F0 variations, sound intensity, speech rate, and timbre information (Banse & Scherer 1996; Juslin & Laukka

---

[1]Department of Otorhinolaryngology/Head and Neck Surgery, University Medical Center Groningen (UMCG), University of Groningen, Groningen, the Netherlands; [2]Pento Speech and Hearing Centers, Apeldoorn, the Netherlands; [3]Research School of Behavioral and Cognitive Neuroscience, Graduate School of Medical Sciences, University of Groningen, Groningen, the Netherlands; [4]Department of Audiology, Faculty of Health Sciences, Hacettepe University, Ankara, Turkey; [5]CNRS UMR 5292, Lyon Neuroscience Research Center, Auditory Cognition and Psychoacoustics, Inserm UMRS 1028, Université Claude Bernard Lyon 1, Université de Lyon, Lyon, France; [6]Department of Audiology and Health Innovation, Research and Development, Sonova AG, Stäfa, Switzerland; [7]Department of Speech-Language Pathology, University of Toronto, Toronto, Ontario, Canada; [8]Department of Psychology, Toronto Metropolitan University, Toronto, Canada; and [9]Auditory Prostheses & Perception Laboratory, Center for Hearing Research, Boys Town National Research Hospital, Omaha, Nebraska, USA.

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and text of this article on the journal's Web site (www.ear-hearing.com).

<zdoi; 10.1097/AUD.0000000000001645>

2001). Recently, von Eiff et al. (2022) manipulated F0 and timbre (aperiodicity, formant frequencies and bandwidths, and spectral tilt) through a voice morphing technique and showed that NH listeners use both F0 and timbre information to similar degrees for vocal emotion recognition. In severe hearing loss and where hearing aid (HA) amplification of acoustic signals is not sufficient, cochlear implants (CIs) replace acoustic hearing with electric stimulation of the auditory nerve to provide hearing. Pediatric CI candidacy criteria are bilateral profound sensorineural hearing loss for children between 9 and 24 months and bilateral severe to profound hearing loss for children older than 2 years, as well as a limited benefit from appropriately fitted HAs (Anne et al. 2022). However, due to factors related to electric stimulation and the resulting channel interactions, signals transmitted are reduced in both spectrotemporal details and dynamic range (for an overview, see Başkent et al. 2016). This likely affects emotional prosodic cues related to spectrotemporal and intensity-related aspects like it is the case for musical emotion perception (Harding et al. 2023), but not speech rate per se, and hence vocal emotion recognition in CI children and adults (Luo et al. 2007; Hopyan-Misakyan et al. 2009; Volkova et al. 2013; Chatterjee et al. 2015, 2023; Everhardt et al. 2020).

HAs amplify the acoustic signal to compensate for elevated hearing thresholds. They also provide other signal processing features to compensate for suprathreshold changes caused by hearing loss. These suprathreshold changes may include reduced spectral and temporal resolution (Van Tasell 1993; Moore 1995; Başkent 2006; Reed et al. 2009; Souza et al. 2015; Brennan et al. 2018), and reduced sensitivity to temporal fine structure (Moore 2008; Hopkins & Moore 2011; Halliday et al. 2019). Loudness recruitment and reduced dynamic range can occur as the maximum tolerable loudness stays the same while the hearing thresholds are higher (Van Tasell 1993; Buus & Florentine 2002). HA amplification alone cannot compensate for such changes and therefore many additional features are provided, such as dynamic range compression to reduce the wide range of the signal amplitude to match the reduced dynamic range (Davidson & Skinner 2006; Pittman et al. 2014; Souza 2016). In high-frequency hearing loss, the inaudible speech sounds can be lowered from high-frequency to low-frequency regions to make them audible (Souza et al. 2013). Although the effects of altered emotion cues can be acute, due to the changes in the perceived signal, they can also be cumulative over time. Exposure to social interactions, factors associated with HA selection and fitting, and rehabilitation from the early years may influence overall HA outcomes in later years (Moeller & Tomblin 2015; McCreery & Walker 2022).

HAs are designed to optimize speech perception and listening comfort (Ching et al. 2001 ; Launer et al. 2016; Pavlovic 1988; Tomblin et al. 2015; Tomblin et al. 2020). Outcome measures with HAs are usually based on speech transmission quality and speech intelligibility (Steeneken & Houtgast 1980; Humes et al. 1986; Holube & Kollmeier 1996; Arehart et al. 2013), and notably, there is currently no standard by which to assess vocal emotion abilities. However, a complete overview of vocal emotion recognition in children with HAs is important as the choices in treatment and management of hearing loss in clinical practice may have consequences for children, particularly in social, cognitive, and academic development (Mauk and Mauk 1992; Hopyan-Misakyan et al. 2009; Roland et al. 2016). Previous study on vocal emotion recognition in child and adult HA users has shown somewhat mixed results. Christensen et al. (2019) tested HA users between 21 and 75 years of age, as well as age-matched normal-hearing listeners, to show that vocal emotion perception was independently hindered both by age and hearing loss. Furthermore, findings from other studies have suggested that a HA benefit may be specific to speech perception and does not transfer to emotion perception. Goy et al. (2018) found no difference in vocal emotion recognition in experienced HA users when comparing aided and unaided listening conditions, while they did see an improvement in speech perception when listeners used their HAs. In addition, Singh et al. (2019) conducted a study with HA users and listeners with hearing loss who did not use HA and found no differences in emotion recognition between the groups. It is unclear whether the reported lack of HA benefit is due to an insufficient transmission of emotion-related acoustic cues by HAs or because these HA users may not be able to correctly map the available cues onto emotional representations. It is important to note that the studies by Goy et al. and Singh et al. included older adult listeners and previous study has shown that emotion recognition is subject to aging effects in addition to effects of hearing loss (Christensen et al. 2019). Considering these different factors, the effect of HAs on emotion recognition may be different for adults and children. Indeed, one relevant study on this topic in a child population of nineteen participants with mild-to-moderate hearing loss, including listeners with and without HAs (8 to 14 years; Cannon & Chatterjee 2019), did not report a significant difference in vocal emotion recognition between children with hearing loss and NH children, using both infant-directed and adult-directed speech. The pure-tone averages across 500, 1000, 2000, and 4000 Hz (PTA4) of the better ear of the children with hearing loss ranged from 13.8 to 48.8 dB HL and 3 of the children with hearing loss did not use a HA in daily life or during testing. Cannon and Chatterjee showed that vocal emotion recognition improved with increasing age and this developmental effect did not differ between children with and without hearing loss. The authors suggest that children with hearing loss may rely on different mechanisms for vocal emotion recognition than NH children due to the finding that in the former group, accuracy scores of older children were associated with their vocabulary level. On the other hand, two other studies on vocal emotion recognition in children (Most & Michaelis 2012) and adolescents (Most & Aviner 2009) showed that HA users had lower recognition scores than NH listeners, but did not differ significantly from CI users. These studies included HA users between the ages of 4 and 6 and between 10 and 15, respectively, but did not investigate developmental effects within these groups.

Previous studies on vocal emotion recognition have looked into the influence of listeners' hearing thresholds, but the results have been inconclusive. Christensen et al. (2019) reported a marginally significant effect of low-frequency hearing thresholds on vocal emotion recognition accuracy in adult HA users. Hearing thresholds at low frequencies are indeed expected to be relevant for vocal emotion recognition as this relates to voice pitch perception (Babaoğlu et al. 2024). As a result, elevated low-frequency hearing thresholds could affect the ability to perceive pitch fluctuations conveying emotional prosody. In addition, other studies have suggested a role for high-frequency hearing in emotion recognition as this may be the range in which some spectral changes related to emotion expressions take place. Some indication for the relevance of hearing thresholds at different frequency regions could be found in a study by Buono et al. (2021) who showed that listeners rated emotional non-speech sounds as less extreme in terms of valence and arousal when stimuli were high-pass filtered or low-pass filtered. This finding supports the hypothesis that reduced audibility at low and/or high frequencies affects emotional ratings of sounds. On the other hand, a number of studies have not found a relation between hearing thresholds and vocal emotion recognition in adults with age-NH or only mild hearing loss (Orbelo et al. 2005; Mitchell 2007; Lambrecht et al. 2012; Dupuis & Pichora-Fuller 2015). It therefore remains unclear whether hearing thresholds would be associated with vocal emotion recognition in children with a wider range of hearing loss degrees and who are using HAs.

Because voice pitch is a robust component of vocal emotion (Banse & Scherer 1996; Juslin & Laukka 2003; Lima & Castro 2011; Lausen & Hammerschmidt 2020), one might expect that participants with greater voice pitch sensitivity would be able to better recognize vocal emotions. Such a link has previously been reported by Globerson et al. (2013) who found that performance in a task focusing on the direction of a pitch change, was a significant predictor of vocal emotion recognition. Other studies have also shown that listeners use vocal tract length (VTL) cues (Chuenwattanapranithi et al. 2009), formant information (Banse & Scherer 1996), or more broadly, voice timbre cues for vocal emotion perception (von Eiff et al. 2022). In addition, the production of emotional speech has been related to differences in vocal tract shaping (Kim et al. 2020), such that listeners with good VTL discrimination may also be better in discriminating emotions.

In this study, we set out to assess how age influences vocal emotion perception in sentence materials for children with HAs, compared with age-matched NH peers. For a comprehensive overview, we measured vocal emotion recognition of school-age children with bilateral HAs who have a wide range of hearing loss degrees and configurations. Because vocal emotion recognition develops during childhood, the quantification of the degree to which a child can perceive vocal emotions is only possible if it is known what recognition level could be expected for their age on the specific test administered. For this purpose and to provide a complete developmental trajectory, we also measured vocal emotion recognition in children and young adults with no hearing loss. Finally, as a first attempt into investigation of any additional factors, we investigated correlations with hearing thresholds in high- and low-frequency regions and with sensitivity measures for F0 and apparent VTL, as measured by Babaoğlu et al. (2024) in the same population as the present study.

MATERIALS AND METHODS

The study presented here is part of a larger project on voice and speech perception in Turkish children and adults: Perception of Indexical Cues in Kids and Adults in Turkish (PICKA-tr). The PICKA-tr project consists of four experiments, previously developed for Dutch (PICKA) and subsequently adapted to Turkish at the University Medical Center Groningen, to assess voice cue discrimination, voice gender categorization, speech perception in the presence of a single-talker speech masker, and vocal emotion recognition (Nagels et al. 2020a, b, 2021, 2024). Participants completed the four experiments in a single test session in the order presented earlier.

Here, we report the results from the vocal emotion recognition test for Turkish children with HAs and children and adults with NH (EmoHI test, previously used in Dutch and English children and adults with NH and in Dutch children with CIs; Nagels et al. 2020b). Participants completed a forced-choice emotion categorization task with meaningless sentences conveying happiness, sadness, and anger. We also further explored correlations of various auditory measures with vocal emotions. First, we assessed correlations with unaided and aided audiometric thresholds in different frequency regions, such as PTA4 (the PTA covering mid-range frequencies between 500 and 4000 Hz, largely overlapping with important speech frequencies and formant regions). In addition, we considered high-frequency PTA (6000 and 8000 Hz; HFPTA), low-frequency PTA (250 and 500 Hz; LFPTA), and the extended low-frequency PTA (125 and 250 Hz; ELFPTA), covering various frequency ranges that may be relevant for vocal emotion perception and following previous studies (Flaherty et al. 2021; Babaoğlu et al. 2024). Second, we assessed correlations with the just-noticeable differences (JNDs) for F0 and VTL from the first test of the PICKA-tr project, tested with the same participants of the present study and previously reported by Babaoğlu et al. (2024).

This study was approved by The Clinical Research Ethical Committee of the university 2019/07-22 (KA19038). Additional permissions were obtained from the Ministry of Education to test NH children at specific schools and from the Ministry of Health to collect data at hospitals and clinics.

Participants

A total of 55 HA children (range = 5.4 to 17.8 years, mean = 10.3 years, SD = 3.5 years), 86 NH children (range = 6.0 to 17.1 years, mean = 11.0 years, SD = 2.9 years), and 68 NH adults (range = 19.1 to 35.0 years, mean = 24.9 years, SD = 4.6 years) took part in the study. Participants with NH were recruited from the general public, via primary schools, and from the local university student population. HA participants were recruited via the University Audiology Clinic, private HA shops, and rehabilitation centers. In the overall study, we included healthy participants who were native speakers of Turkish, had normal or corrected-to-normal vision, and who had no history of neurological, developmental, motor, or language disorders, based on self-reports for adults and older children, or parental reporting for younger children. All children attended state schools, thus receiving similar levels of education. Furthermore, NH and HA children showed similarities regarding socioeconomic background and music education. None of the children received any additional music education other than what is included in the school program (approximately 1 hour/week). Maternal education level differed somewhat within and between the groups of NH and HA children (for NH children, university: N = 60, high school: N = 23, primary school: N = 2, unknown: N = 1; for HA children, university: N = 8, high school: N = 10, primary school: N = 32, unknown: N = 5). For the NH groups, we only included participants with hearing thresholds ≤20 dB HL at audiometric frequencies of 0.5, 1, 2, and 4 kHz and with a negative

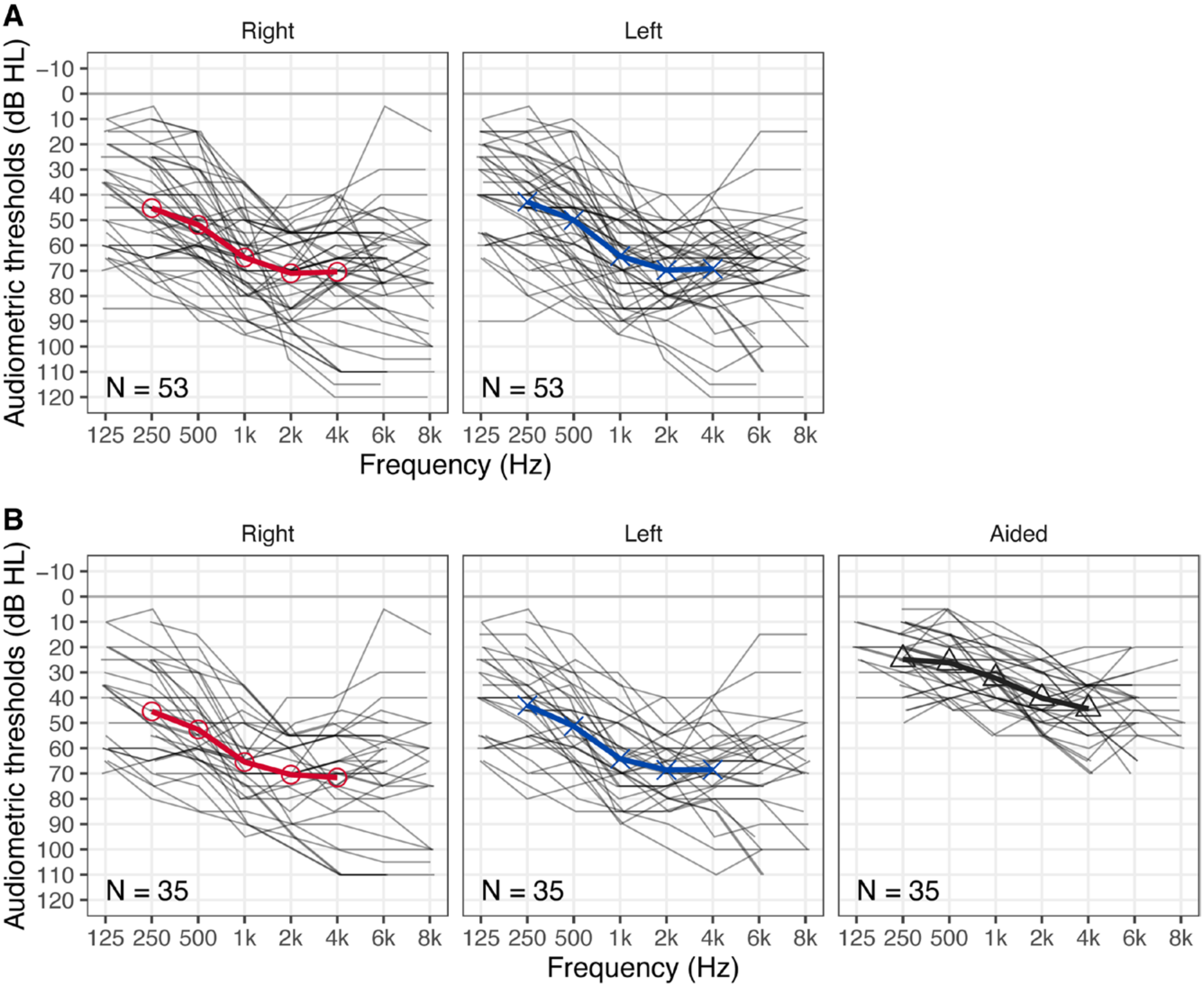

Fig. 1. Unaided and aided audiometric thresholds for the hearing-aided children from whom this information was available from medical records. A, Unaided pure-tone audiograms of hearing-aided children, available from N = 53 children, shown for right and left ears in left and right panels, respectively, and for the widest range of audiometric frequencies available for each child. B, Unaided (left and middle panels) and aided (right panel) pure-tone audiograms, available from N = 35 hearing-aided children. Light gray lines represent the individual hearing thresholds across the audiometric frequencies between 125 Hz and 8 kHz, while bold blue, red, and black lines indicate the group means for right, left, or both ears, across the audiometric frequencies between 250 and 4 kHz.

history of unresolved middle ear issues. For the group of HA users, we recruited children who had used bilateral HAs for at least 6 months at the time of testing.

Hearing-Aided Participants

All HA children who participated in this study were healthy bilateral HA users for a minimum of 6 months (HA use duration range = 0.6 to 14.0 years, mean = 6.2 years, SD = 3.6 years). Hearing loss etiology included congenital and acquired hearing loss, and ototoxicity for 1 participant, based on parental reporting. However, not all parents were able to provide this information. Furthermore, none of the HA children had any major problems in speech and language development, other than what could be expected as a result of their hearing loss, as reported by their parents in the demographic questionnaire. The degree of sensorineural hearing loss of the participants was determined by PTA4, the pure-tone air conduction thresholds averaged across 500, 1000, 2000, and 4000 Hz for the better ear, following the classification system of the World Health Organization for hearing impairment (World Health Organization 2021). PTA4 range of the HA children was 37.5 to 97.5 dB HL, including moderate (N = 11), moderately severe (N = 23), severe (N = 15), profound (N = 4), and unknown (N = 2) hearing loss. In addition, we had access to aided audiometric thresholds for a subset of the HA children (N = 35). Figure 1 shows the individual and averaged unaided (row A) and unaided and aided (row B) audiometric thresholds. All children used bilateral Phonak hearing aids (Phonak, Aurora, IL; Sky: N = 38; Naida: N = 10; Bolero: N = 4; Audeo: N = 3). Children predominantly used behind-the-ear type HAs (N = 52), with few using a receiver-in-canal type (N = 3). Earpieces were mostly custom-made earmolds with standard tubes, followed by domes with slim tubes. Fitting formulas and data logging information of HAs were checked through the Phonak Target fitting software (version 7.1., 2021; Sonova AG, Stäfa Switzerland). Almost all children were fitted with the DSL-V5 pediatric prescriptive formulation (DSL-V5: N = 50; NAL-NL2: N = 1; Adaptive Phonak: N = 1; unknown N = 3).

TABLE 1. Speaker demographics and their voice characteristics

Speaker	Age (yr)	Gender	Height (m)	Mean F0 (Hz)	F0 Range (Hz)
T2	36	F	1.68	302	201–437
T3	27	M	1.85	167	101–296
T5	25	F	1.63	283	199–429
T6	24	M	1.75	168	87–286

For 53 of the 53 HA children, parents were able to report the age at the start of HA use. Almost all children (N = 52) had started using their HAs before the age of 10, with an average of 3.6 years (range = 0.25 to 10.0 years, SD = 2.5 years), and 1 participant used their HAs starting from age 16.0. HA use duration ranged from 0.9 to 14.6 years (mean = 6.4 years, SD = 3.6 years). Data logging from 28 children also showed that the daily HA use for 26 of these children was at least 7 hours (range = 7.4 to 15.8 hours, mean = 12.0 hours, SD = 2.5 hours). The remaining 2 children were identified as outliers with little daily device use (2.3 and 0.4 hours). Their records showed that one of these participants experienced battery problems during the period in which data logging took place. The other participant made little use of their HA due to relatively good hearing thresholds in the lower frequency range (10 to 25 dB HL at frequencies up to 1 kHz).

Stimuli

The two sentences used in this study, *Koun se mina lod belam* [kʌun sə mina: lɔd be:lam] and *Nekal ibam soud molen* [ne:kal ibam sʌut mo:lən], were originally derived from the Geneva Multimodal Emotion Portrayal Corpus materials (Bänziger et al. 2012). These pseudospeech sentences consist of plausible syllables, but without meaning in any language, making the stimuli suitable for use across different languages and study populations. Utterances of these sentences were produced by 4 adult native Dutch speakers (2 female, 2 male) expressing three basic emotions (happiness, sadness, and anger). Speaker demographics and their voice characteristics are presented in Table 1. Further details of the sentence recordings are described by Nagels et al. (2020b), and the complete stimulus set is available online*. The speech recordings presented in this study consisted of 3 productions of each emotion by each speaker (3 utterances × 3 emotions × 4 speakers), resulting in 36 stimuli in total. Four additional stimuli (one each for happy and angry, and two for sad) were used for four practice trials before the actual test block. Each practice stimulus was produced by 1 of the 4 speakers that produced the test stimuli, but the practice stimuli were not part of the experimental stimulus set. All participants were presented with the same four practice stimuli to familiarize them with the test interface, with the pseudospeech sentences, and with the voices of each of the four speakers.

Experimental Setup

For the NH group, hearing thresholds were screened with an Interacoustics AS608B portable screening audiometer (Interacoustics, Middelfart, Denmark) and RadioEar DD45 headphones (RadioEar, Middelfart, Denmark). The vocal emotion recognition test interface was displayed on a Lenovo Yoga touchscreen laptop. A custom MATLAB (R2018b) script provided experimental control for the test (The MathWorks, Inc., Natick, MA, USA). The stimuli were presented through Sennheiser HD 380 Pro headphones for NH participants and through Logitech Z200 speakers, placed approximately 70 cm from the participant, for HA children. For both the headphones and speakers, the stimulus presentation level was calibrated to 65 dB SPL with a sound level meter (Svantek 979; Svantek Sp, Warsaw, Poland) and Kemar simulator (45BB KEMAR Head and Torso; G.R.A.S. Sound & Vibration A/S, Holte, Denmark). During the experiment, participants were seated in a relatively quiet room at the testing site, such as the library in the schools or in a sound booth in clinics or HA shops. HA children completed the experiment with their own bilateral HAs using their daily device settings.

Procedure

All participants and, when applicable, their parents or legal guardians were given detailed information about the study and provided written informed consent before the experiment. NH children and adults were screened at octave frequencies between 500 and 4000 Hz to verify their hearing status. Adult participants and parents of participating children filled out a demographic questionnaire, which took 3 to 4 minutes to complete. This questionnaire contained questions about the participant's age, education level, musical training, language and speech development, general health, hearing health history, and hearing status. For HA children, information about hearing status and HA use was provided by additional questionnaire items and parental consultation. Parents of HA children also completed the Children's Alexithymia Measure (CAM) as a short and reliable measurement of children's ability to identify and describe feelings (Way et al. 2010). The CAM consists of 14 questions that are scored between 0 and 3, and parents took approximately 2 to 3 minutes to complete the questionnaire. Results from the CAM will be reported elsewhere.

The vocal emotion recognition test, the EmoHI test (Nagels et al. 2020b), was conducted using the child-friendly game-like interface for both child and adult participants (Fig. 2). The test started with a practice session of four trials. After the practice session with 4 trials, the data collection started in which all 36 test items were presented in randomized order in a single block. During the experiment, a parrot and three clowns with happy, angry, and sad facial expressions were presented on the screen. Participants were instructed to listen to the sentence uttered by the parrot, and then identify the vocally expressed emotion by tapping on the clown with the matching facial expression on the touchscreen. The order of the three clowns was randomized across participants. Feedback was provided through falling confetti in case of a correct answer, or through the parrot shaking its head

---

*The EmoHI corpus is available from Zenodo, http://doi.org/10.5281/zenodo.7997063.

Fig. 2. The child-friendly game-like interface of the vocal emotion recognition test. The illustrations were made by Jop Luberti for the study by Nagels et al. (2020b), and published under the CC BY NC 4.0 license (https://creativecommons.org/licenses/by-nc/4.0/).

in case of an incorrect answer. Participants were informed of the testing duration through the other two clowns in the background climbing the ladder after every two trials. The end of the experiment was indicated by the second clown jumping in the pool. Participants took approximately 5 to 7 minutes to complete the vocal emotion recognition test and the testing time for the complete PICKA-tr test battery was approximately 50 to 60 minutes.

Data Analysis

Accuracy scores were calculated for overall vocal emotion recognition. Hit rates and false alarms were used to determine participants' sensitivity index *d*' for each emotion category (Macmillan & Creeman 2005). To calculate *d*', the difference between z-transformed hit rates and false alarms for each emotion category was divided by √2. To avoid infinite z-transformed values, all hit rates or false-alarm rates of 0 or 1 were corrected by half a trial. The *d*' values were averaged across emotions for visualization and for the estimation of the distribution of NH children and adults in the first analysis. During the study, it became apparent that a hidden option was activated that allowed participants to skip to the next trial without giving a response. This accidentally happened on two occasions, once for a HA child and once for an adult with NH. For these participants, 35 instead of 36 trials were considered in the data analysis. All statistical analyses were performed in R (version 4.2.3; R Core Team 2020).

First, we performed a quantile regression analysis of the data of NH children and adults, to model the age-typical development and to investigate where children with HAs would fall in this distribution, using the qgam package (version 1.3.4; Fasiolo et al. 2021) in R. The quantile regression analysis is based on generalized additive models (GAMs) and allows for an estimation of the distribution of response accuracy scores as a function of age by modeling different quantiles (e.g., percentiles). Quantile regression does not assume a parametric distribution and therefore does not have the assumption of normality.

In a second analysis, we performed a group analysis using GAMs to compare vocal emotion recognition expressed as d' between NH and HA children. Note that NH adults were not included in this analysis. Individual *d*' values were entered into the model, using mgcv (version 1.8.42; Wood 2003, 2004, 2011, 2017; Wood et al. 2016), itsadug (version 2.4.1; Rij et al. 2022), and gratia (version 0.8.1; Simpson 2023) packages in R. To examine the development of accuracy scores for each group as a function of age, we used the following model:

$$d \sim \text{group} + s\,(\text{age},\ \text{by} = \text{group},\ \text{bs} = ''\ \text{cs}'')$$

The spline used for age was a cubic regression spline, with shrinkage, fitted per participant group (NH, HA), with the k-parameter set to 5. The dependent variable was *d*', averaged across the three emotion categories. All fittings were done using the restricted maximum likelihood (ML) method. To perform the same analysis per emotion, the "group" factor was replaced with the interaction between group and emotion, and the age splines were fitted per group and per emotion.

Finally, to assess whether other predictors could account for the remaining variability not explained by age or hearing group, we compared the model above to a similar model where the tested predictor was added. We used various PTAs as predictors. In addition to PTA4, PTAs at HFPTA and low frequencies (LFPTA, ELFPTA) were determined for both unaided and aided listening conditions. The comparison of two models

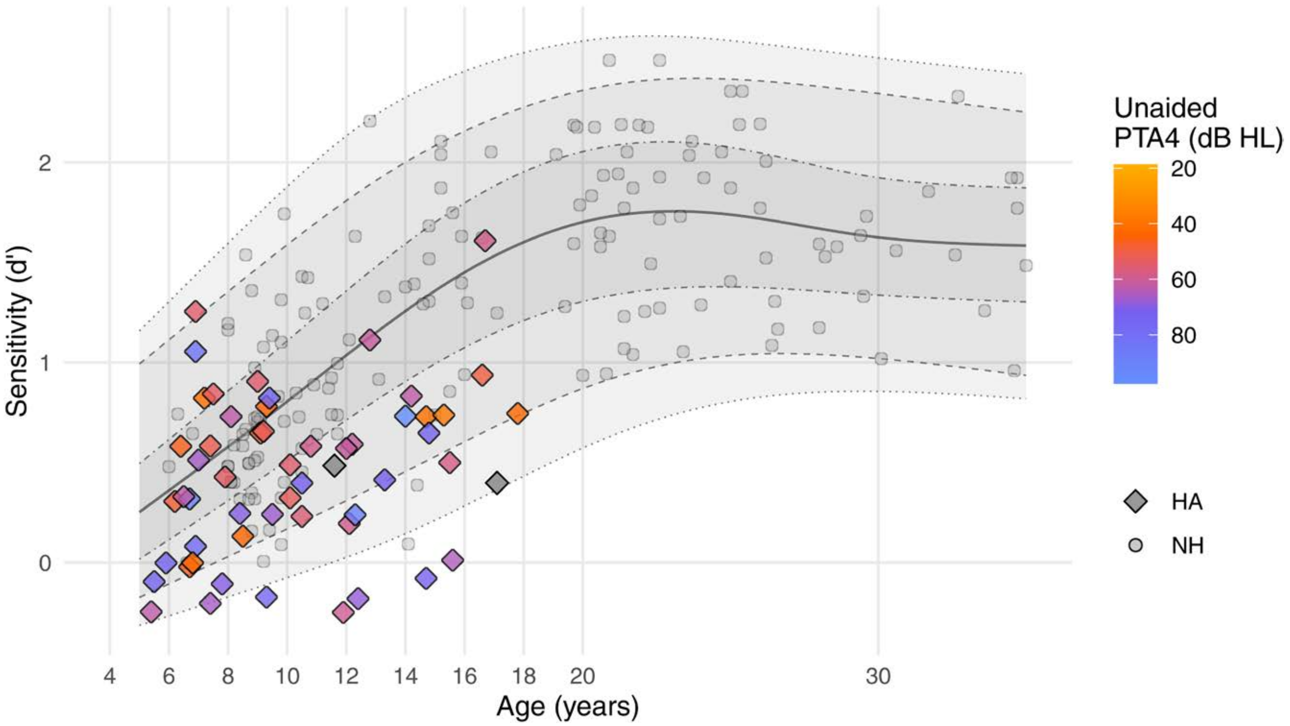

Fig. 3. The distribution for age-typical development based on *d′* values of children and adults with NH shown with the individual *d′* values of the participants with NH (gray circles, N = 154) or HAs (colored and gray diamonds, N = 55) superimposed. The age-typical distribution was determined via the quantile regression analysis of *d′* values of NH children and adults, based on a generalized additive model with a cubic regression spline with shrinkage, and as a function of age. Shaded areas represent the 1st, 5th, 25th, 50th (median, in bold gray line), 75th, 95th, and 99th percentiles. The color of the diamonds represents the unaided PTA4 for the HA children (N = 53) and the gray diamonds represent the HA children (N = 2) for whom the unaided PTA4 was not available. HA indicates hearing aids; NH, normal hearing; PTA, pure-tone average.

yields a χ^2 and a p value. For these comparisons, the models had to be (re-)fitted with the unrestricted ML method, rather than the restricted ML, because the latter implies a rescaling of the data that depends on the parametric factors present. The comparison of these two models gives us a conservative estimate of the influence of the predictor on the data independently from the effect of age. Furthermore, we assessed whether the sensitivity to voice cues is associated with vocal emotion recognition. To this aim, we used F0 and VTL JNDs of the same participants as previously reported by Babaoğlu et al. (2024). Because age is a common predictor for emotion recognition and voice discrimination, we correlated the two factors expressed as quantiles of the NH distribution over age. The NH distribution for vocal emotion recognition was estimated in the first analysis described earlier and the NH distribution for voice discrimination was estimated as described by Babaoğlu et al.

RESULTS

Figure 3 shows the quantile regression of the sensitivity index *d*', averaged across the three emotions. The distribution of *d*' values of the NH children and adults was estimated as a function of age, and shown in gray shading. Individual data of the children and adults with NH are plotted as gray circles, while the data of children with HAs are plotted as colored diamonds. Both NH and HA groups show a large variation at all ages. A quantile regression on the raw accuracy scores of NH children and adults was also conducted and is presented in Supplementary Figure 1, Supplemental Digital Content, http://links.lww.com/EANDH/B607. Accuracy scores of NH adults ranged from 64% to 97%, with *d*' ranging from 0.94 to 2.51.

The vocal emotion recognition of the NH participants seems to plateau around 20 years old. From the age of 6 to the plateau, the median accuracy scores steadily increase from around 40%-correct to about 80%-correct, and from 0.25 to 1.76 in the case of *d*' values. In comparison, although some HA children seem to perform on par with their peers, a number of them seem to demonstrate limited sensitivity. Overall, 12 of the 55 HA children (22%) had *d*' values above the median of the NH distribution. More strikingly, 21 of the children with HAs (38%) had *d*' values below the 10th percentile of the NH distribution. Table 2 provides a complete overview of the number of HA children with *d*' values below and above various percentiles of the NH distribution.

TABLE 2. Number of HA children with *d*' values below and above various percentiles based on the NH distribution

Percentile	HA Children Scoring Below (%)	HA Children Scoring Above (%)
1st	6 (11)	49 (89)
5th	12 (22)	43 (78)
10th	21 (38)	34 (62)
25th	36 (65)	19 (35)
50th	43 (78)	12 (22)
75th	51 (93)	4 (7)
90th	53 (96)	2 (4)
95th	54 (98)	1 (2)
99th	55 (100)	0 (0)

HA, hearing aids; NH, normal hearing.

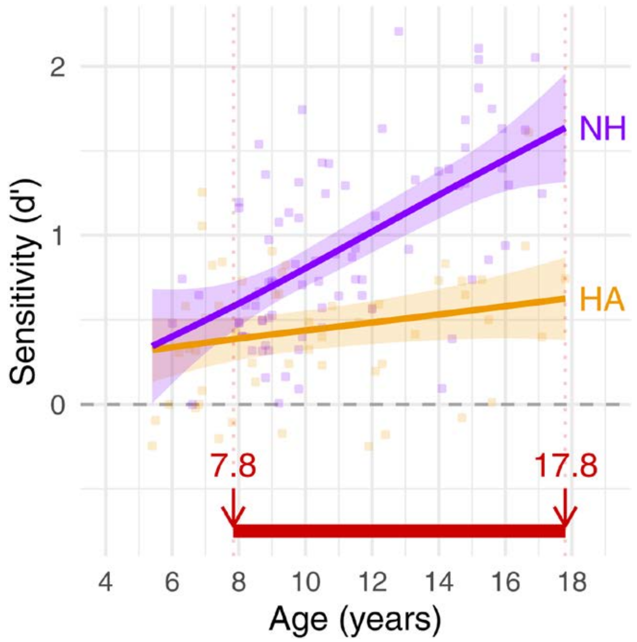

Fig. 4. GAM analysis for the comparison of overall vocal emotion recognition (sensitivity, expressed as *d′*) averaged across emotions, as a function of age for the NH (bold purple line) and HA (bold orange line) groups, restricted to the child participants. The shaded areas surrounding the regression lines represent the 95% credibility interval, while the circles show the individual scores. The red bar indicates the age interval over which the spline fittings for the two groups are significantly different from each other. The horizontal dashed line represents chance level performance ($d' = 0$). GAM indicates generalized additive model; HA, hearing aids; NH, normal hearing.

Inspecting the sensitivity index of individual participants per group, vocal emotion recognition development as a function of age per group seems to differ between the NH and HA groups. To assess this question, we fitted GAM regressions as a function of age per group (Fig. 4), limited to children only. The analysis showed that performance from both NH [$F(1.621, 136.525) = 12.05$, $p < 0.001$] and HA [$F(0.855, 136.525) = 0.77$, $p < 0.05$] children significantly depended on age. However, visual inspection of Figure 4 and the *F* values shown earlier indicate that the development of *d*' values of the HA children appears to be slower than that of the NH participants. When comparing the fits for each group, it appears that the two groups differ significantly from age 8 onwards.

TABLE 3. Effect of adding various unaided and aided pure-tone audiometric thresholds, averaged across high frequencies (6000 and 8000 Hz; HFPTA), mid frequencies (500, 1000, 2000, and 4000 Hz; PTA4), low frequencies (250 and 500 Hz; LFPTA), and extended low frequencies (125 and 250 Hz; ELFPTA), to the GAM model for the HA children

Unaided	HFPTA	ΔAIC = 1.74, $\chi^2(1.00) = 1.876$, $p = 0.05$, N = 53
	PTA4	ΔAIC = 0.34, $\chi^2(1.00) = 1.194$, $p = 0.12$, N = 53
	LFPTA	ΔAIC = –1.75, $\chi^2(1.00) = 0.165$, $p = 0.57$, N = 53
	ELFPTA	ΔAIC = –0.69, $\chi^2(1.00) = 0.649$, $p = 0.26$, N = 35
Aided	HFPTA	ΔAIC = –0.39, $\chi^2(1.00) = 0.804$, $p = 0.21$, N = 27
	PTA4	ΔAIC = –1.97, $\chi^2(1.00) =0.004$, $p = 0.93$, N = 35
	LFPTA	ΔAIC = –1.99, $\chi^2(1.00) = 0.002$, $p = 0.96$, N = 32
	ELFPTA	ΔAIC = –0.67, $\chi^2(1.00) = 0.664$, $p = 0.25$, N = 10

The p values are not corrected for multiple comparisons.

ΔAIC, Akaike information criterion difference; ELFPTA, extended low-frequency PTA; HA, hearing aids; HFPTA, high-frequency PTA; LFPTA, low-frequency PTA; PTA, pure-tone average.

We have also analyzed the sensitivity per emotion category. Figure 5 shows the *d*' values per group as a function of age, broken down per emotion category in the different panels. It appears that sensitivity for "sad" seems to be driving most of the age effect in HA children [$F(1.314, 1123.375) = 41.09$; $p < 0.001$], while sensitivity for "angry" and "happy" was not significantly dependent on age ($p > 0.6$). For NH children, the age effect was significant for all emotion categories (*p*s < 0.001).

Correlation With Audiometric Thresholds

Table 3 reports the effects of adding the various unaided and aided PTAs to the GAM regression, in addition to age. We found that none of the PTAs had convincing predicting value for vocal emotion recognition. It is important to note that all akaike information criterion differences were small, indicating that any contribution of the PTAs was small.

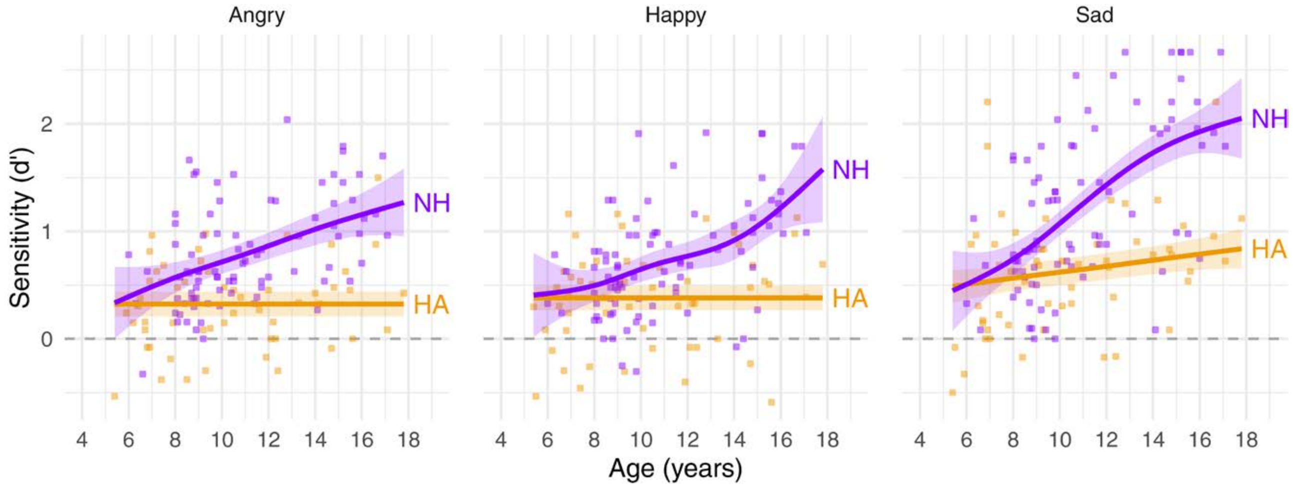

Fig. 5. Same as Figure 4, except that *d′* is shown per emotion category per group. HA indicates hearing aids; NH, normal hearing.

Correlation With Voice Cue Discrimination

To assess whether vocal emotion recognition may be associated with F0 and VTL sensitivity in HA children, we performed two correlations between vocal emotion recognition and either F0 or VTL JNDs, both expressed as quantiles of the NH distribution. Neither F0 JNDs [$r^2 < 0.001$, $t(53) = 0.19$, $p = 0.85$] nor VTL JNDs [$r^2 < 0.001$, $t(53) = 0.03$, $p = 0.98$] showed any sign of correlation with vocal emotion recognition based on *d*' values. Indeed, visual inspection of the raster plots in Figure 6 indicates that some participants who are performing better than the median of the NH group for vocal emotion recognition (right half of the panels) may also, at the same time, be showing worse JNDs than the NH participants (upper half of the panels).

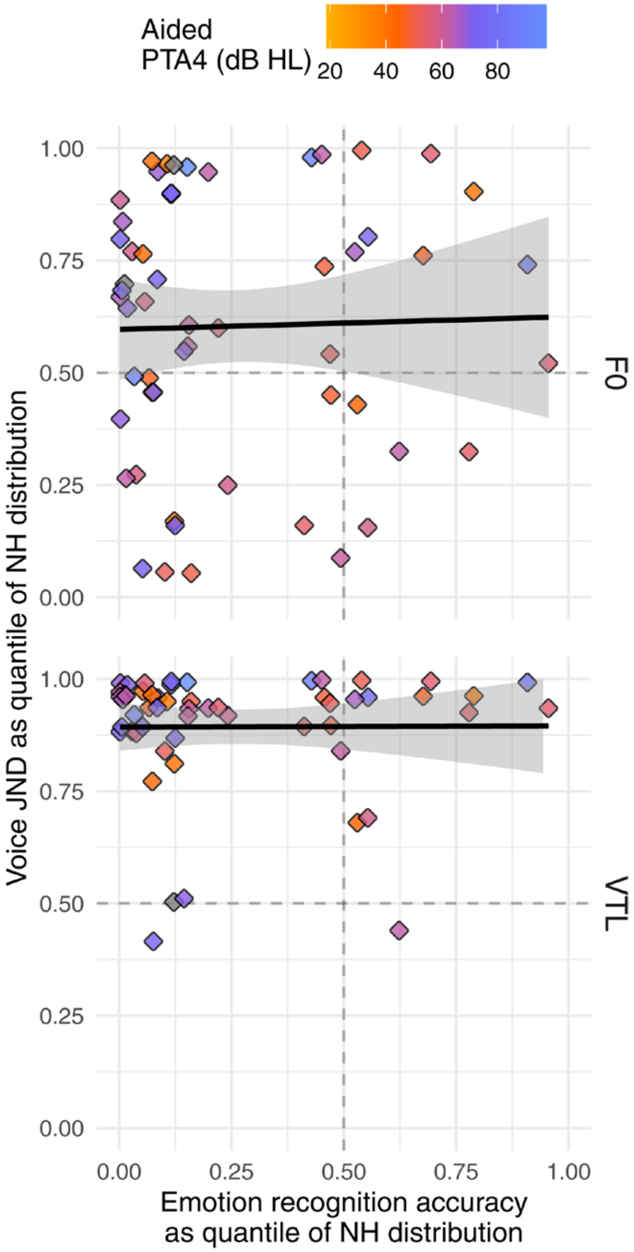

Fig. 6. The relation between vocal emotion recognition and voice cue discrimination. Voice just-noticeable differences for fundamental frequency (F0; upper panel) and VTL (lower panel) from Babaoğlu et al. (2024) are shown as a function of vocal emotion recognition accuracy, both expressed as quantiles of the NH distribution (Fig. 3). The individual scores are the same as in Figure 3, with the same color coding. NH indicates normal hearing; PTA, pure-tone average; VTL, vocal tract length.

DISCUSSION

In this study, we provide a comprehensive overview of vocal emotion recognition in school-age children who have a wide range of hearing losses and who use bilateral HAs. We additionally provide baseline developmental data from NH children and adults, so that changes due to hearing loss could be evaluated with respect to what could be expected from age-typical levels. Vocal emotion recognition was tested for three basic emotions: happiness, sadness, and anger. We found a developmental effect for NH children and adults, with overall vocal emotion recognition improving with increasing age, and reaching a plateau around age 20. In HA children, vocal emotion recognition also improved with age, albeit differently from the NH group from around age 8 onwards. Data from individual children and per emotion category provided more nuances about the potential overlap or difference between the two groups. Inspection of emotion categories showed no developmental effect for HA children for the emotions of angry and happy. Inspection of individual data indicated that a number of HA children, some even with severe hearing loss, did perform above the median of the baseline performance of NH children. On the other hand, a number of HA children scored lower than age-expected levels, with some even at chance levels. Finally, we found no evidence that audiometric thresholds could explain the variability in emotion recognition in HA children, nor did we find a correlation between emotion recognition performance and sensitivity to F0 or VTL voice cues.

The developmental trajectories in NH listeners found in this study are in line with previous studies (Aguert et al. 2013; Sauter et al. 2013; Chronaki et al. 2015; Grosbras et al. 2018; Nagels et al. 2020b; Amorim et al. 2021; Filippa et al. 2022), showing that emotion recognition improves during childhood and adolescence as a function of age and over many years. In terms of the use of the relatively newly developed EmoHI test, this study complements the study of Nagels et al. (2020b) by using the same vocal emotion recognition test in a new cohort of NH children and adults, with an extended age range. Having used the same paradigm as Nagels et al. (2020b) , it should be noted that the overall scores of the Turkish NH listeners showed more variability and were generally lower in the present study than in that of Nagels et al. (2020b) who tested Dutch NH listeners. Although the variability of scores in the present study (range = 64% to 97%) was larger compared with Nagels et al. (2020b) , it is comparable to the ranges reported in previous studies with NH adult listeners (~48% to 100%, Globerson et al. 2013; Christensen et al. 2019; ~60% to 92%; Amorim et al. 2021; ~50% to 80%). Furthermore, even though the EmoHI test was developed to be language-independent by using meaningless pseudosentences, the stimuli were recorded from Dutch talkers. Cultural influences in the way emotions are expressed or perceived could therefore affect the recognizability of emotions in different populations. An in-group advantage for vocal emotion recognition has been demonstrated across various cultures (for reviews, see Elfenbein & Ambady 2002; Laukka & Elfenbein 2021). This advantage could be driven by production differences across cultures, but recognizability of the vocal cues expressing emotions could also be affected by familiarity with the language that carries these expressions (Nakai et al. 2023). As a result, not only the origin of the talker could have affected the lower performance in the present study, but also the fact that the pseudosentences used in this study may be phonologically more related to the Dutch than the Turkish language, given that they were originally chosen to represent plausible phoneme combinations for a number of Western languages (Bänziger et al. 2012). These potential effects of phonological familiarity (Fleming et al. 2014) are also highlighted by the finding that English adults performed quite similarly to Dutch adults with the same EmoHI test materials (Nagels et al. 2020b). It is therefore possible that the Turkish HA children may have had a greater disadvantage than Turkish NH children when faced with even a slightly foreign intonation or prosodic pattern. Notwithstanding these possible cultural effects, a comparison between NH and HA groups with the same linguistic background still revealed developmental effects, similar to the results from Dutch and English populations tested by Nagels et al. (2020b). The present study thus shows that the EmoHI test can also be used to examine developmental trajectories in Turkish listeners, despite the linguistic, and potentially also cultural, differences in how emotions are communicated. Here, we included NH participants in the full age range between 5 and 35 years and found that vocal emotion recognition reached a plateau around the age of 20. Similar to the present study, earlier work by Grosbras et al. (2018) and Amorim et al. (2021) also examined nonlinear effects of age in vocal emotion recognition in children and adolescents and found that adult-level performance was reached between 14 and 15 years of age and around the age of 20, respectively. This consistency across the studies is relatively surprising given the differences across the study methods and participant cohorts. Different studies mentioned earlier included different age ranges in the respective studies (up to age 17 years in Grosbras et al. (2018) and up to age 35 years in the present study), which may affect the level of uncertainty about the exact slope of vocal emotion recognition accuracy as a function of age. Furthermore, there were methodological differences, such as the type of stimuli and the number of emotion categories presented (4 and 10 nonverbal affective vocalizations in Amorim et al. (2021) and Grosbras et al. (2018), respectively, versus 3 in our study). The consistent finding of similar and long developmental trajectories indicates that overall, vocal emotion recognition likely relies also on general cognitive and developmental mechanisms, such as correct interpretation and labeling of emotion categories, and not only hearing the specific acoustic cues of the emotion categories (Albanese et al. 2010). For example, previous research suggested that children's verbal abilities could be related to understanding emotions, as language could help both naming the emotions and also learning the complexity required for understanding others' emotions (von Salisch et al. 2013). Hence, despite the differing methodological approaches and potentially differing linguistic and cultural differences of the cohorts, it is evident from the current and previous studies that vocal emotion recognition develops over many years during childhood, even in NH. Given the long development period, any potential difficulties for vocal emotion recognition in HA children can therefore only be identified with respect to what would be realistic based on their age.

Like NH children, HA children also showed a significant developmental effect, albeit seemingly less pronounced than NH children. A development of vocal emotion recognition in HA children has been shown in one earlier study by Cannon and Chatterjee (2019) who included 8- to 14-year-old children with mild-to-moderate hearing loss. The present study extends this finding by showing that age-related improvement in vocal emotion recognition can be detected in a larger sample with a slightly wider age range (5.4 to 17.8 years) and including more severe and a wider range of degrees of hearing loss (moderate to profound). However, both visual inspection of the age effects in Figures 4 and 5, and the GAM regressions point to a stronger age effect in NH children compared with HA children. Although the overall *d*' values of the NH and HA groups did not significantly differ for the youngest children, they did significantly differ from age 8 years. Furthermore, for the three emotion categories, there was a significant age effect for the NH children, while HA children only showed a significant effect of age for sadness. Based on previous findings, it is possible that children may have selected "sad" as a default response and that the lack of an escape option in the current experimental design may have introduced or strengthened a response bias. In line with this possibility, Chronaki et al. (2015) reported a bias for "sad" in NH children within an age range (4 to 11 years) that partially overlaps with the age range of the children in the present study. Inspecting the confusion matrices for our younger groups, however, we do not see concentrated confusions for the sad category for both NH and HA children (Supplementary Figure 1, Supplemental Digital Content, http://links.lww.com/EANDH/B607). Furthermore, the order of the clowns was randomized across participants to avoid a bias due to a preference for selecting, for example, the middle clown when participants did not know what to respond.

A significant group effect between the age-matched NH and HA listeners has also been reported in earlier research with children (Most & Michaelis 2012) and adolescents (Most & Aviner 2009), but was not shown in the study by Cannon and Chatterjee (2019). Like the present study, the studies by Most and Michaelis (2012) and Most and Aviner (2009) included HA children with a wide range of hearing loss degrees, with PTA3 of the better ear ranging from 40 to 115 dB HL and from 73 to 97 dB HL, respectively (versus a PTA4 of the better ear ranging from 37.5 to 97.5 dB HL in the present study). On the other hand, the study by Cannon and Chatterjee included only children with mild-to-moderate hearing loss with better-ear PTA4s between 13.8 and 48.8 dB. As such, it may be the case that this group of children is not as much affected by suprathreshold changes as a result of their hearing loss and the relatively mild levels of amplification and compensation needed from the HAs. Moreover, three participants in the study by Cannon and Chatterjee did not use a HA in daily life or during testing. It is therefore possible that, at the group level, the inclusion of children with more severe degrees of hearing loss could account for the mixed findings of the abovementioned studies. Finally, differences in study design could also have led to the differences regarding the presence or absence of a group effect. Both the present study and the study by Most and Michaelis used meaningless sentences, while Cannon and Chatterjee and Most and Aviner used meaningful and semantically neutral sentences. Previously, Geers et al. (2013) reported that the processing of linguistic information (i.e., "what is said") and indexical information (talker-specific information about e.g., emotional state or personality traits) is significantly correlated in children with CIs. Together with the findings of Cannon and Chatterjee that linguistic ability plays a role in vocal emotion recognition by children with hearing loss, these results could suggest that the use of meaningless versus meaningful sentences may have an impact on HA children's ability to recognize emotions in speech. The lower emotion recognition scores for HA children in the study by Most and Aviner, who did find a group difference between NH and HA children using meaningful sentences, could then be due to the number of emotion categories that were presented since the probability to give an incorrect response increases when more emotion categories, and thus response options, are used. Most and Aviner used six emotion categories, Cannon and Chatterjee used five, Most and Michaelis used four, and the present study used three categories. The combined effects of number of categories and sentence content may therefore lead to differences in task difficulty. Despite these differences, results of the present study are generally consistent with previous findings from cohorts with similar degrees of hearing loss, showing a difference in vocal emotion recognition between NH and HA children at the group level.

Although as a group, HA children seem to experience more difficulties in vocal emotion recognition compared with NH children, shown by significant group differences in ages 8 years and older, Figure 3 and Table 2 indicate that a considerable proportion of HA children had *d*' values comparable to those of NH children. About one-third of the HA children (19 of 55; 35%) had *d*' values above the 25th percentile of the NH distribution. Despite the group difference, a more nuanced approach of these results therefore brings up the possibility that not all HA users are challenged in their perception of vocal emotions, and more certainly, not to the same degree. Hence, even though HAs have been developed and optimized for speech perception (Pavlovic 1988; Ching et al. 2001; Tomblin et al. 2015, 2020; Launer et al. 2016), data from this study indicate that they can also support age-typical development of vocal emotion recognition. In addition, it is possible that the group difference is due to a delayed development of HA children rather than an overall deficit in vocal emotion perception. Longitudinal follow-up research could further explore this possibility by measuring developmental trajectories of individuals from childhood up to adulthood. These possibilities also provide a more optimistic perspective on the added value of HAs than the conclusions by Goy et al. (2018) and Singh et al. (2019) do. In these two studies, older adult HA users performed a vocal emotion recognition test with and without wearing their HAs. There was no significant difference in emotion recognition accuracy scores between the two listening modes even though the participants from the study by Goy et al. did show an improvement in word recognition in the aided compared with the unaided listening condition. However, these studies presented their results at the group level only, such that it remains possible that some of their individual participants also experienced a HA benefit for vocal emotion recognition. A considerable overlap in vocal emotion recognition scores between NH and HA listeners can also be seen in the individual data from Christensen et al. (2019), while at the group level, HA adults showed lower vocal emotion recognition scores than age-matched NH adults. The potentially differing interpretations of the same data based on group versus individual data analyses suggest that we need to be cautious of the implications of our findings for clinical practice. The consideration of individual data is clinically relevant, as it provides valuable information that can be used to identify those HA users that may benefit from additional rehabilitation targeting emotion perception, and also points to the urgency for vocal emotion audiometry to be included in clinical screenings.

Because of this study's inclusive approach, four children with profound hearing loss participated. These children were 6.7, 6.9, 12.3, and 14.0 years old and had PTA4s of 90.0, 83.8, 90.0, and 97.5 dB HL, respectively, implying that these children could meet CI candidacy criteria. Although individuals with such severe degrees of hearing loss may face unique challenges in perceiving vocal emotions, we decided to include these participants to have a realistic representation of pediatric HA users. If there are HA children in the general population without appropriately fitted devices or with too little or too profound hearing loss for HA use, this would also be reflected in our study population. For this study, we do not have information on whether these 4 participants with profound hearing loss have considered receiving CIs, but from the available data logging, we derived a daily device use of more than 7 hours per day for the 6.9-year-old participant and more than 13 hours per day for the 14.0-year-old participant. In addition, based on the aided and unaided PTA4, the two children with an unaided PTA4 of 90.0 dB HL showed a substantial improvement of 46.3 and 60.0 dB HL when using their HAs. Finally, while the performance of these children showed some variation and was rather low for three of the four children (*d*' of 0.32, 1.05, 0.24, and 0.73), their data did not stand out as outliers compared with the overall NH distribution as a function of age (Fig. 3). The *d*' values of the 2 younger HA children, aged 6.7 and 6.9, was within the distribution of NH listeners of the same age (above the 25th and 75th percentile, respectively). For these children, it is therefore still possible that vocal emotion perception will improve with continued HA use. The *d*' values of the 2 older HA children, aged 12.3 and 14.0 years, was on the low side for their age group (above the 1st and 5th percentile, respectively), but there were a number of other HA children of the same age and with less severe hearing loss that had lower *d*' values, indicating that challenges in vocal emotion perception are not unique to those listeners with the most profound levels of hearing loss. Together, these data show that, even with profound hearing loss, some of these children were able to recognize vocal emotions to the degree that could be expected based on the distribution of their NH peers.

The overlap of vocal emotion recognition scores of individual NH and HA children is also due to a large variability in both groups. Such variability is likely a result of emotion recognition in general being linked with other factors than acute perception of affective acoustic cues only, as even in NH children, a relation to cognitive and linguistic factors had been shown (Griffiths et al. 2020; Schlegel et al. 2020). In HA children, these factors would also play a role. Indeed, nonverbal cognition, as measured by visual pattern replication and completion tasks, has previously been shown to be related to emotion recognition in school-age CI users (Chatterjee et al. 2023). This predictor interacted with hearing age, indicating that especially the younger children with higher nonverbal cognition showed higher emotion recognition scores. However, the functional mechanism underlying the association between nonverbal cognition and emotion recognition is not clear and could be addressed in future research. Furthermore, the lack of semantic context in this test may make it more difficult for some children than for others in this period in which children are still developing. (Morton and Trehub 2001) have shown that children between 4 and 10 years of age develop from relying more on linguistic content to relying more on prosodic information when recognizing emotions in speech. In addition to this development, if HA children have reduced access to acoustic cues conveying emotions, it is possible they rely even more on the linguistic information. Although it was not possible to address these factors with our current experimental design, new test materials may help assess these hypotheses. Finally, the audio materials used in this study were recorded to convey emotions in a subtle way and the use of multiple talkers may have resulted in larger variations in the way emotions were vocally expressed. These factors are also likely to accentuate the variability in the results from both children and adults.

In addition to cognitive factors, long- and short-term access to the affective acoustic cues play a crucial role in vocal emotion recognition in listeners with hearing loss. It is possible that differences in accumulated auditory experience account for some of the variability in HA children. The importance of auditory access for the development of communication skills is delineated in the cumulative auditory experience model. This framework describes the importance of early auditory exposure and access to language interactions for the development of language and executive function skills in children with hearing loss (Moeller & Tomblin 2015; McCreery & Walker 2022). In addition, the use of mental state language when interacting with toddlers has been related to children's emotion understanding abilities (Taumoepeau & Ruffman 2006), and children with moderate hearing loss have been shown to be at a disadvantage compared with their peers with NH in the development of empathy, putting them at risk for social-emotional difficulties (Dirks et al. 2017). Because vocal emotion recognition does not only rely on the access to or the perception of relevant acoustic cues but also on the ability to interpret these cues and associate them with the correct emotion category or label, sufficient exposure to emotions and opportunities of incidental learning is needed to be able to learn these categories. Correctly categorizing the presented emotion therefore reflects a late decision-related process of emotion recognition, after perceiving and interpreting the auditory input, and hearing loss may have both acute and prolonged effects on this process. Although the specific role of early linguistic input in vocal emotion recognition has not yet been investigated, it is very well possible that differences in early auditory exposure may explain some of the variability in emotion recognition scores from this study. A recent study on psychosocial difficulties in children with hearing loss indirectly supports this idea by showing that psychosocial difficulties and vocal emotion perception are associated with general communication skills (de Jong et al. 2023).

Many children who participated in this study had a high chance of receiving a HA early in life thanks to the neonatal screening program in Turkey that was implemented in the early 2000s and has been the nationwide standard screening program since 2008. Nevertheless, this study cohort included children with a wide range of initial HA fits (age 0.25 to 16.0 years) and we cannot rule out the possibility that some children who started using their HAs at older ages had levels of hearing loss that were undetected or not sufficiently treated during a prolonged period. Cases of undetected or untreated hearing loss may also accentuate differences in auditory exposure as mentioned earlier and consequently contribute to the variability in scores of HA children. To account for this source of variability, future studies could be specifically designed to take the time between a hearing loss diagnosis and initial HA fit, as well as the average daily HA use, into consideration.

With both acute effects of hearing loss on access to relevant acoustic cues and the potential cumulative effects of hearing loss on vocal emotion recognition development in mind, the variability seen in HA children was further examined by considering several factors that may explain the individual differences in the results. Some factors that could affect auditory input did not vary much within the group of HA children. The demographic information indicated that all HA children in our study were healthy, received good audiological care and seemingly from a relatively young age, used well-fitted earmolds and HAs, and had been using their HAs for at least 6 months or longer and for many hours per day. Finally, there was little variability among the HA children when considering several education-related factors. In particular, all children attended state schools providing similar forms of education and received similar hours of musical education at school. One factor that demonstrated some variability both within the group of HA children and between the NH and HA children at the group level is maternal education. Maternal education of the participants included university, high school, as well as primary school level. For HA children, a relatively large proportion reported maternal education to be at primary school level, whereas for NH children, maternal education was mostly reported to be at the university level. These differences are possibly the result of different recruitment strategies for NH and HA participants. HA children were typically recruited through clinical care centers, whereas NH children were mostly recruited through the co-authors' personal and professional networks. As such, maternal education, a contributing factor of the home learning environment, could account for some of the observed group differences in vocal emotion recognition (Li et al. 2023). Other factors that could play a role but we do not have any information on are the quality and quantity of language interactions of the HA children during their early childhood. We therefore cannot rule out that some of the variability seen in our data stems from other cumulative factors. Ideally, longitudinal studies starting during infancy and covering both the quantity and quality of linguistic and social interactions would be of great value to better understand the effects of early auditory exposure on later emotion recognition abilities.

The effects of acute access to affective acoustic cues on vocal emotion recognition were further assessed by considering differences in hearing loss degree and configuration within our cohort. The present study was designed for maximum inclusion such that all healthy children who used bilateral HAs for at least 6 months and who were able to perform the experiment could participate. As a result, the HA group included children with a wide age range (5.4 to 17.8 years) and with degrees of hearing loss ranging from moderate to profound (unaided PTA4: 37.5 to 97.5 dB HL). It can be expected that for many of the HA children in this study, vocal emotion recognition is affected by suprathreshold effects in addition to the effects of elevated hearing thresholds. For example, spectral resolution can be compromised for moderate to severe hearing losses (Rosen et al.

1990; Nelson 1991; Baker & Rosen 2002; Başkent 2006). In line with this, in the same group of HA children of this study, Babaoğlu et al. (2024) previously reported a reduced sensitivity to VTL compared with age-matched NH children. Since VTL perception is related to spectral resolution of the broad bandwidth of the speech signal, these results provide an indication that some of the HA children in this study may have wider auditory filters, which HAs seem to not entirely be able to compensate for. On the other hand, F0 JNDs of these HA children showed a large overlap with those of NH children, especially at older ages (Babaoğlu et al. 2024). In that study, variability in F0 sensitivity was shown to be related to hearing thresholds in low-frequency ranges that overlap with the average F0 range of the talkers that produced the stimulus material. Together, these observations suggest that variations in the degree and configuration of hearing loss of HA children, along with accompanying potential suprathreshold deficiencies, could have both acute and long-term effects and can lead to differences in both voice cue perception, and also vocal emotion recognition.

To further examine the effects of hearing loss degree and configuration, we added various measures of unaided hearing thresholds as predictors to our statistical models. For a subset (N = 35) of these HA children we had access to audiometry data measured with their HAs. For this subgroup, we performed further analyses by adding aided hearing thresholds to our statistical models. In addition to the aided and unaided PTA4s, we also assessed aided and unaided HFPTA, LFPTA, and ELFPTA as predictors. Because vocal emotion recognition typically requires good pitch perception, LFPTA and ELFPTA were added as model predictors to capture hearing thresholds that cover the F0 range of the speakers in the emotion test (87 to 437 Hz; see Table 1 and Nagels et al. 2020b). Our recent study on voice cue sensitivity in the same group of children showed that unaided ELFPTA was a significant predictor for F0 sensitivity (Babaoğlu et al. 2024), suggesting that access to low-frequency information affects voice pitch perception, which may consequently also affect vocal emotion perception. To fully explore the various frequency regions of the audiogram, unaided and aided HFPTA was also added as a predictor for emotion recognition. None of the correlations tested in the present study showed a significant correlation with vocal emotion recognition as expressed as *d*'. We therefore conclude that neither unaided nor aided hearing thresholds have any predictive value for vocal emotion recognition. Previous studies have shown mixed results on the predictive value of hearing thresholds for vocal emotion recognition in HA users. Singh et al. (2019) reported a significant correlation between vocal emotion recognition accuracy and PTA4 averaged across the left and right ear in older adult HA users (>67 years). In a sample of adult HA users with a wider age range (22 to 74 years), Christensen et al. (2019) found that LFPTA had a marginally significant effect on vocal emotion recognition, while PTA4 and HFPTA were not predictive of emotion recognition accuracy. Finally, Most and Michaelis (2012) examined the effect of aided and unaided PTA3 of 500, 1000, and 2000 Hz and of the pure-tone threshold at 500 Hz on vocal emotion recognition in young children (4 to 6 years) with hearing loss and did not find any significant correlation, in line with the findings of the present study. However, when children were presented with video recordings of emotional expressions, performance in the auditory-visual modality was correlated with hearing thresholds at 500 Hz.

In an additional exploratory analysis, we further assessed whether sensitivity to the voice cues of F0 and VTL had any predictive value in vocal emotion recognition performance. Previous literature focusing on CI users suggests that a degradation of F0 cues in the signal transmitted by the CI poses challenges in emotional-prosody perception (Luo et al. 2007; Chatterjee et al. 2015; Everhardt et al. 2020; Nagels et al. 2020b). We performed correlation analyses using JNDs from the same group of participants that were previously reported by Babaoğlu et al. (2024). Neither F0 JNDs nor VTL JNDs correlated with vocal emotion recognition expressed as *d*'. Given the importance of pitch-based cues for vocal emotion recognition, it is somewhat surprising that we did not find a correlation between F0 sensitivity and emotion recognition scores. Babaoğlu et al. previously reported the promising finding that F0 sensitivity of the same sample of HA children improved with increasing age to the extent that the JNDs of HA children were similar to that of NH children during teenage years. This suggests that despite the effects of hearing loss and likely due to consistent HA use, voice pitch cues are somewhat available to the HA children. However, to correctly categorize vocal emotions, it is not sufficient to perceive the relevant acoustic cues, but these cues need to be correctly mapped to emotion categories. This mapping develops during childhood and could be affected by hearing loss as well. Previous research in NH children demonstrated a dissociation between voice cue sensitivity and the ability to use these voice cues in a voice gender categorization task (Nagels et al. 2020a). It is possible that a similar dissociation can be seen in the present study population for voice cue perception and vocal emotion recognition. For VTL, its relevance for vocal emotion perception is less well-described, but has also been assigned a relevant role in vocal expressions of emotion (Kim et al. 2020). The assessment of VTL JNDs as a predictor for vocal emotion recognition was therefore more exploratory in nature. Still, the lack of correlations between voice cue sensitivity and vocal emotion recognition could also be attributed to the design of the JND measurements. The perception of voice cues in the study by Babaoğlu et al. was assessed by measuring sensitivity to static F0 and VTL cues, which are possibly not appropriate to capture sensitivity to emotion-related acoustic cues that are more dynamic in nature (such as prosodic cues, Frick 1985; Banse & Scherer 1996; Everhardt et al. 2020). However, Globerson et al. (2013) used several pure-tone pitch discrimination tasks and found that only those tasks in which participants had to focus on the direction of the pitch change, rather than merely detecting a change in pitch, were significant predictors of performance in a vocal emotion recognition task. In the JND task from Babaoğlu et al., measuring sensitivity to mean F0, participants were not asked to focus on the direction of the change in F0 but rather to detect any difference (i.e., manipulations of average voice F0) in an odd-one-out task. Neither the JND task of Babaoğlu et al. nor the experiment of (Globerson et al. 2013) thus targeted pitch contours specifically, but the findings by Globerson et al. indicate that other measures of pitch discrimination might still be relevant to follow-up on in future research on vocal emotion perception in HA children. In addition to audiometric and psychoacoustic factors, it is also possible that other variables related to HA fit, HA use, and hearing loss etiology may have an effect on vocal emotion perception. For a subset of HA children for whom the fitting report was available (N = 30), it appeared that vent size varied from open to occluded, with various intermediate sizes reported (0.6–4.5 mm). Vent size may have an influence on vocal emotion perception by affecting the access to low-frequency cues conveying vocal emotions and could be taken into consideration in future studies focusing on the effect of various HA features and fitting procedures, as well as factors related to hearing loss etiology. For further discussion of the possible effects of HA features on voice perception in this study population, see also Babaoğlu et al.

Summarizing our most important findings, our results with relatively large cohorts show that investigating individual children's data in relation to what would be expected for their age is crucial to achieve a realistic assessment for their vocal emotion recognition. The proportion of children who show age-typical sensitivity to vocal emotions indicates HA benefits for vocal emotion perception. The HA children who did not score at age-typical levels likely could benefit from additional and specialized rehabilitation and tools provided for caregivers, and to identify them, a child-appropriate and validated diagnostic test for vocal emotion recognition could be a helpful addition to standard audiological care. With a more complete picture of the abilities and difficulties of children with HAs, one can then think of potential approaches to further support communication abilities, including vocal emotion perception, in children with hearing loss. In a recent study with young HA children (4 to 9 years), Yeshoda et al. 2020 have provided a first indication that vocal emotion recognition training may improve emotional-prosody perception. Other research on CI users has also alluded that music-based training may improve vocal emotion perception in both children (Good et al. 2017) and adults (Fuller et al. 2018). Future research could therefore focus on further investigating the long-term effects of different training programs on vocal emotion recognition in children and assess for which children such training could be especially valuable. Given the developmental trajectories observed also in HA children, with good support, these HA children may have a good chance to further improve their vocal emotion recognition.

ACKNOWLEDGMENTS

The authors are grateful to all children, parents, students, and colleagues who participated in this study, the Turkish Ministry of National Education (T.C. Milli Eğitim Bakanlığı) for general support, and the Hacettepe University audiology clinic, state schools, hearing aid shops, and rehabilitation centers for their support with the recruitment of participants. The authors thank Paolo Toffanin, Leanne Nagels, Iris van Bommel, Evelien Birza, Jacqueline Libert, Jemima Phillpot, Marta Matos Lopes, and Jop Luberti (illustrations) for their contribution to the development of the child-friendly game-like interface, the authors' children and children's friends for their help with piloting and improving the initial versions of the test, Jennifer Breetveld, Wijnna Beuving, Gesina Posthumus-Ottema, Esther Steenbergen for research support, and Christina Elsenga for help in data sharing.

This work was funded by Sonova Industry Support, a VICI Grant (No. 918-17-603) from the Netherlands Organization for Scientific Research (NWO) and the Netherlands Organization for Health Research and Development (ZonMw), and Heinsius Houbolt Funds. The study was performed within the framework of the Laboratoire d'Excellence Centre Lyonnais d'Acoustique (ANR-10-LABX-0060) of Université de Lyon within the program "Investissements d'Avenir" (ANR-16-IDEX-0005) operated by the French National Research Agency (ANR).

E.G., L.N., and D.B. designed the experiment. G.B., B.O.Y., P.E., F.U., M.C., E.Y., G.S. contributed to parts of the design related to participant inclusion and demographics. G.B., B.O.Y., P.E., P.D., F.U., D.B. evaluated the hearing aid settings. G.B., B.O.Y., and P.E. collected the data. L.R., G.B., B.O.Y., P.E., E.G., and D.B. reviewed the data. L.R. and E.G. analyzed the data with help from G.B., B.O.Y., P.E., and D.B.; M.C. gave additional advice on data analysis. L.R. and D.B. wrote the manuscript with contributions from G.B., B.O.Y., P.E., and E.G. All authors discussed the results, their interpretation and implications, and commented on the manuscript at all stages.

Consent for the research was provided by the Clinical Research Ethical Committee of Hacettepe University 2019/07-22 (KA19038).

The authors have no conflicts of interest to disclose.

Address for correspondence: Laura Rachman, Department of Otorhinolaryngology/Head and Neck Surgery, University Medical Center Groningen, Hanzeplein 1 9713 GZ, Groningen, the Netherlands. E-mail: l.rachman@rug.nl

OPEN PRACTICES

This manuscript qualifies for an Open Data badge and an Open Materials badge. The materials and data have been made publically available at https://doi.org/10.34894/MEZP1E and https://zenodo.org/records/7997063. More information about the Open Practices Badges can be found at https://journals.lww.com/ear-hearing/pages/default.aspx.

Received March 20, 2024; accepted January 15, 2025

REFERENCES

Aguert, M., Laval, V., Lacroix, A., Gil, S., Le Bigot, L. (2013). Inferring emotions from speech prosody: Not so easy at age five. *PLoS One*, *8*, e83657.

Albanese, O., De Stasio, S., Di Chiacchio, C., Fiorilli, C., Pons, F. (2010). Emotion comprehension: The impact of nonverbal intelligence. *J Genet Psychol*, *171*, 101–115.

Amorim, M., Anikin, A., Mendes, A. J., Lima, C. F., Kotz, S. A., Pinheiro, A. P. (2021). Changes in vocal emotion recognition across the life span. *Emotion*, *21*, 315–325.

Anne, S., Brown, K. D., Goldberg, D. M., Adunka, O. F., Kenna, M., Chien, W., Buchman, C. A. (2022). Pediatric bilateral sensorineural hearing loss: Minimum test battery and referral criteria for cochlear implant candidacy evaluation. *Otolaryngol Head Neck Surg*, *166*, 405–409.

Arehart, K. H., Souza, P., Baca, R., Kates, J. M. (2013). Working memory, age and hearing loss: Susceptibility to hearing aid distortion. *Ear Hear*, *34*, 251–260.

Babaoğlu, G., Rachman, L., Ertürk, P., Özkişi Yazgan, B., Sennaroğlu, G., Gaudrain, E., Başkent, D. (2024). Perception of voice cues in school-age children with hearing aids. *J Acoust Soc Am*, *155*, 722–741.

Baker, R. J., & Rosen, S. (2002). Auditory filter nonlinearity in mild/moderate hearing impairment. *J Acoust Soc Am*, *111*, 1330–1339.

Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *J Pers Soc Psychol*, *70*, 614–636.

Bänziger, T., Mortillaro, M., Scherer, K. R. (2012). Introducing the Geneva Multimodal expression corpus for experimental research on emotion perception. *Emotion*, *12*, 1161–1179.

Başkent, D. (2006). Speech recognition in normal hearing and sensorineural hearing loss as a function of the number of spectral channels. *J Acoust Soc Am*, *120*, 2908–2925.

Başkent, D., Gaudrain, E., Tamati, T. N., Wagner, A. (2016). Perception and psychoacoustics of speech in cochlear implant users. In A. T. Carcace, E. de Kleine, A. G. Hort P. van Dijk (Eds.) *Scientific Foundations of Audiology: Perspectives From Physics, Biology, Modeling, and Medicine* (pp. 285–319). Plural Publishing, Inc.

Blasi, A., Mercure, E., Lloyd-Fox, S., Thomson, A., Brammer, M., Sauter, D., Deeley, Q., Barker, G. J., Renvall, V., Deoni, S., Gasston, D., Williams, S. C. R., Johnson, M. H., Simmons, A., Murphy, D. G. M. (2011). Early specialization for voice and emotion processing in the infant brain. *Curr Biol*, *21*, 1220–1224.

Brennan, M. A., McCreery, R. W., Buss, E., Jesteadt, W. (2018). The influence of hearing-aid gain on gap-detection thresholds for children and adults with hearing loss. *Ear Hear*, *39*, 969–979.

Buono, G. H., Crukley, J., Hornsby, B. W. Y., Picou, E. M. (2021). Loss of high- or low-frequency audibility can partially explain effects of hearing loss on emotional responses to non-speech sounds. *Hear Res*, *401*, 108153.

Buus, S., & Florentine, M. (2002). Growth of loudness in listeners with cochlear hearing losses: Recruitment reconsidered. *J Assoc Res Otolaryngol*, *3*, 120–139.

Cannon, S. A., & Chatterjee, M. (2019). Voice emotion recognition by children with mild-to-moderate hearing loss. *Ear Hear*, *40*, 477–492.

Chatterjee, M., Gajre, S., Kulkarni, A. M., Barrett, K. C., Limb, C. J. (2023). Predictors of emotional prosody identification by school-age children with cochlear implants and their peers with normal hearing. *Ear Hear*, *45*, 411–424.

Chatterjee, M., Zion, D. J., Deroche, M. L., Burianek, B. A., Limb, C. J., Goren, A. P., Kulkarni, A. M., Christensen, J. A. (2015). Voice emotion recognition by cochlear-implanted children and their normally-hearing peers. *Hear Res*, *322*, 151–162.

Cheng, Y., Lee, S. -Y., Chen, H. -Y., Wang, P. -Y., Decety, J. (2012). Voice and emotion processing in the human neonatal brain. *J Cogn Neurosci*, *24*, 1411–1419.

Ching, T. Y. C., Dillon, H., Katsch, R., Byrne, D. (2001). Maximizing Effective Audibility in Hearing Aid Fitting. *Ear Hear*, *22*, 212–224.

Christensen, J. A., Sis, J., Kulkarni, A. M., Chatterjee, M. (2019). Effects of age and hearing loss on the recognition of emotions in speech. *Ear Hear*, *40*, 1069–1083.

Chronaki, G., Hadwin, J. A., Garner, M., Maurage, P., Sonuga-Barke, E. J. S. (2015). The development of emotion recognition from facial expressions and non-linguistic vocalizations during childhood. *Br J Dev Psychol*, *33*, 218–236.

Chuenwattanapranithi, S., Xu, Y., Thipakorn, B., Maneewongvatana, S. (2009). Encoding emotions in speech with the size code: A perceptual investigation. *Phonetica*, *65*, 210–230.

Davidson, L. S., & Skinner, M. W. (2006). Audibility and speech perception of children using wide dynamic range compression hearing aids. *Am J Audiol*, *15*, 141–153.

de Jong, T. J., van der Schroeff, M. P., Stapersma, L., Vroegop, J. L. (2023). A systematic review on the impact of auditory functioning and language proficiency on psychosocial difficulties in children and adolescents with hearing loss. *Int J Audiol*, *63*, 675–685.

Dirks, E., Ketelaar, L., van der Zee, R., Netten, A. P., Frijns, J. H. M., Rieffe, C. (2017). Concern for others: A study on empathy in toddlers with moderate hearing loss. *J Deaf Stud Deaf Educ*, *22*, 178–186.

Dupuis, K., & Pichora-Fuller, M. K. (2015). Aging affects identification of vocal emotions in semantically neutral sentences. *J Speech Lang Hear Res*, *58*, 1061–1076.

Elfenbein, H. A., & Ambady, N. (2002). On the universality and cultural specificity of emotion recognition: A meta-analysis. *Psychol Bull*, *128*, 203–235.

Everhardt, M. K., Sarampalis, A., Coler, M., Başkent, D., Lowie, W. (2020). Meta-analysis on the perception of linguistic and emotional prosody in cochlear implant users and vocoder simulations. *Ear Hear*, *41*, 1092–1102.

Fasiolo, M., Wood, S. N., Zaffran, M., Nedellec, R., Goude, Y. (2021). qgam: Bayesian Nonparametric Quantile Regression Modeling in R. *J Stat Softw*, *100*, 1–31.

Filippa, M., Lima, D., Grandjean, A., Labbé, C., Coll, S. Y., Gentaz, E., Grandjean, D. M. (2022). Emotional prosody recognition enhances and progressively complexifies from childhood to adolescence. *Sci Rep*, *12*, Article 1.

Flaherty, M. M., Buss, E., Leibold, L. J. (2021). Independent and combined effects of fundamental frequency and vocal tract length differences for school-age children's sentence recognition in a two-talker masker. *J Speech Lang Hear Res*, *64*, 206–217.

Fleming, D., Giordano, B. L., Caldara, R., Belin, P. (2014). A language-familiarity effect for speaker discrimination without comprehension. *Proc Natl Acad Sci USA*, *111*, 13795–13798.

Frick, R. W. (1985). Communicating emotion: The role of prosodic features. *Psychol Bull*, *97*, 412–429.

Fuller, C. D., Galvin, J. J., Maat, B., Başkent, D., Free, R. H. (2018). Comparison of two music training approaches on music and speech perception in cochlear implant users. *Trends Hear*, *22*, 2331216518765379.

Geers, A. E., Davidson, L. S., Uchanski, R. M., Nicholas, J. G. (2013). Interdependence of linguistic and indexical speech perception skills in school-age children with early cochlear implantation. *Ear Hear*, *34*, 562–574.

Globerson, E., Amir, N., Golan, O., Kishon-Rabin, L., Lavidor, M. (2013). Psychoacoustic abilities as predictors of vocal emotion recognition. *Atten Percept Psychophys*, *75*, 1799–1810.

Good, A., Gordon, K. A., Papsin, B. C., Nespoli, G., Hopyan, T., Peretz, I., Russo, F. A. (2017). Benefits of music training for perception of emotional speech prosody in deaf children with cochlear implants. *Ear Hear*, *38*, 455–464.

Goy, H., Pichora-Fuller, M. K., Singh, G., Russo, F. A. (2018). Hearing aids benefit recognition of words in emotional speech but not emotion identification. *Trends Hear*, *22*, 1–16.

Griffiths, S., Goh, S. K. Y., Norbury, C. F.; the SCALES team. (2020). Early language competence, but not general cognitive ability, predicts children's recognition of emotion from facial and vocal cues. *PeerJ*, *8*, e9118.

Grosbras, M. -H., Ross, P. D., Belin, P. (2018). Categorical emotion recognition from voice improves during childhood and adolescence. *Sci Rep*, *8*, 14791.

Halliday, L. F., Rosen, S., Tuomainen, O., Calcus, A. (2019). Impaired frequency selectivity and sensitivity to temporal fine structure, but not envelope cues, in children with mild-to-moderate sensorineural hearing loss. *J Acoust Soc Am*, *146*, 4299–4314.

Harding, E. E., Gaudrain, E., Hrycyk, I. J., Harris, R. L., Tillmann, B., Maat, B., Free, R. H., Başkent, D. (2023). Musical emotion categorization with vocoders of varying temporal and spectral content. *Trends Hear*, *27*, 1–19.

Holube, I., & Kollmeier, B. (1996). Speech intelligibility prediction in hearing-impaired listeners based on a psychoacoustically motivated perception model. *J Acoust Soc Am*, *100*, 1703–1716.

Hopkins, K., & Moore, B. C. J. (2011). The effects of age and cochlear hearing loss on temporal fine structure sensitivity, frequency selectivity, and speech reception in noise. *J Acoust Soc Am*, *130*, 334–349.

Hopyan-Misakyan, T. M., Gordon, K. A., Dennis, M., Papsin, B. C. (2009). Recognition of affective speech prosody and facial affect in deaf children with unilateral right cochlear implants. *Child Neuropsychol*, *15*, 136–146.

Humes, L. E., Dirks, D. D., Bell, T. S., Ahlstbom, C., Kincaid, G. E. (1986). Application of the articulation index and the speech transmission index to the recognition of speech by normal-hearing and hearing-impaired listeners. *J Speech Lang Hear Res*, *29*, 447–462.

Juslin, P. N., & Laukka, P. (2001). Impact of intended emotion intensity on cue utilization and decoding accuracy in vocal expression of emotion. *Emotion*, *1*, 381–412.

Juslin, P. N., & Laukka, P. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychol Bull*, *129*, 770–814.

Kim, J., Toutios, A., Lee, S., Narayanan, S. S. (2020). Vocal tract shaping of emotional speech. *Comput Speech Lang*, *64*, 101100.

Lambrecht, L., Kreifelts, B., Wildgruber, D. (2012). Age-related decrease in recognition of emotional facial and prosodic expressions. *Emotion*, *12*, 529–539.

Launer, S., Zakis, J. A., Moore, B. C. J. (2016). Hearing Aid Signal Processing. *Hearing Aids*, *56*, 93–130.

Laukka, P., & Elfenbein, H. A. (2021). Cross-cultural emotion recognition and in-group advantage in vocal expression: A meta-analysis. *Emot Rev*, *13*, 3–11.

Lausen, A., & Hammerschmidt, K. (2020). Emotion recognition and confidence ratings predicted by vocal stimulus type and prosodic parameters. *Humanit Soc Sci Commun*, *7*, 1–17.

Li, S., Tang, Y., Zheng, Y. (2023). How the home learning environment contributes to children's social–emotional competence: A moderated mediation model. *Front Psychol*, *14*, 1–19.

Lima, C. F., & Castro, S. L. (2011). Speaking to the trained ear: Musical expertise enhances the recognition of emotions in speech prosody. *Emotion*, *11*, 1021–1031.

Luo, X., Fu, Q. -J., Galvin, J. J. (2007). Cochlear implants special issue article: Vocal emotion recognition by normal-hearing listeners and cochlear implant users. *Trends Amplif*, *11*, 301–315.

Macmillan, N. A., & Creelman, C. D. (2005). *Detection Theory: A User's Guide* (2nd ed.). Lawrence Erlbaum Associates Publishers

Mauk, G. W., Mauk, P. P. (1992). Somewhere, Out There: Preschool Children with Hearing Impairment and Learning Disabilities. *Topics Early Childhood Special Edu*, *12*, 174–195.

McCreery, R. W., & Walker, E. A. (2022). Variation in auditory experience affects language and executive function skills in children who are hard of hearing. *Ear Hear*, *43*, 347–360.

Mitchell, R. L. C. (2007). Age-related decline in the ability to decode emotional prosody: Primary or secondary phenomenon? *Cogn Emot*, *21*, 1435–1454.

Moeller, M. P., & Tomblin, J. B. (2015). An introduction to the outcomes of children with hearing loss study. *Ear Hear*, *36*, 4S–13S.

Moore, B. C. J. (1995). *Perceptual Consequences of Cochlear Damage*. Oxford University Press.

Moore, B. C. J. (2008). The role of temporal fine structure processing in pitch perception, masking, and speech perception for normal-hearing and hearing-impaired people. *J Assoc Res Otolaryngol*, *9*, 399–406.

Morton, J. B., Trehub, S. E. (2001). Children's Understanding of Emotion in Speech. *Child Development*, *72*, 834–843.

Most, T., & Aviner, C. (2009). Auditory, visual, and auditory-visual perception of emotions by individuals with cochlear implants, hearing aids, and normal hearing. *J Deaf Stud Deaf Educ*, *14*, 449–464.

Most, T., & Michaelis, H. (2012). Auditory, visual, and auditory–visual perceptions of emotions by young children with hearing loss versus children with normal hearing. *J Speech Lang Hear Res*, *55*, 1148–1162.

Nagels, L., Gaudrain, E., Vickers, D., Hendriks, P., Başkent, D. (2020a). Development of voice perception is dissociated across gender cues in school-age children. *Sci Rep*, *10*, 5074.

Nagels, L., Gaudrain, E., Vickers, D., Matos Lopes, M., Hendriks, P., Başkent, D. (2020b). Development of vocal emotion recognition in school-age children: The EmoHI test for hearing-impaired populations. *PeerJ*, *8*, e8773–e8714.

Nagels, L., Gaudrain, E., Vickers, D., Hendriks, P., Başkent, D. (2021). School-age children benefit from voice gender cue differences for the perception of speech in competing speech. *J Acoust Soc Am*, *149*, 3328–3344.

Nagels, L., Gaudrain, E., Vickers, D., Hendriks, P., Başkent, D. (2024). Prelingually deaf children with cochlear implants show better perception of voice cues and speech in competing speech than postlingually deaf adults with cochlear implants. *Ear Hear*, *45*, 952–968.

Nakai, T., Rachman, L., Sarah, P. A., Okanoya, K., Aucouturier, J. -J. (2023). Algorithmic voice transformations reveal the phonological basis of language-familiarity effects in cross-cultural emotion judgments. *PLoS One*, *18*, e0285028.

Nelson, D. A. (1991). High-level psychophysical tuning curves: Forward masking in normal-hearing and hearing-impaired listeners. *J Speech Hear Res*, *34*, 1233–1249.

Orbelo, D. M., Grim, M. A., Talbott, R. E., Ross, E. D. (2005). Impaired comprehension of affective prosody in elderly subjects is not predicted by age-related hearing loss or age-related cognitive decline. *J Geriatr Psychiatry Neurol*, *18*, 25–32.

Pavlovic, Chaslav (1988). Articulation index predictions of speech intelligibility in **hearing** aid selection. *Asha*, *30*, 63–65.

Picou, E. M., Singh, G., Goy, H., Russo, F., Hickson, L., Oxenham, A. J., Buono, G. H., Ricketts, T. A., Launer, S. (2018). Hearing, emotion, amplification, research, and training workshop: Current understanding of hearing loss and emotion perception and priorities for future research. *Trends Hear*, *22*, 1–24.

Pittman, A. L., Pederson, A. J., Rash, M. A. (2014). Effects of fast, slow, and adaptive amplitude compression on children's and adults' perception of meaningful acoustic information. *J Am Acad Audiol*, *25*, 834–847.

Reed, C. M., Braida, L. D., Zurek, P. M. (2009). Review of the literature on temporal resolution in listeners with cochlear hearing impairment: A critical assessment of the role of suprathreshold deficits. *Trends Amplif*, *13*, 4–43.

Roland, L., Fischer, C., Tran, K., Rachakonda, T., Kallogjeri, D., Lieu, J. E. C. (2016). Quality of Life in Children with Hearing Impairment. *Otolaryngol Head Neck Surg*, *155*, 208–219.

Rosen, S., Faulkner, A., Smith, D. A. J. (1990). The psychoacoustics of profound hearing impairment. *Acta Otolaryngol*, *109*(sup469), 16–22.

Sauter, D. A., Panattoni, C., Happé, F. (2013). Children's recognition of emotions from vocal cues. *Br J Dev Psychol*, *31*, 97–113.

Schlegel, K., Palese, T., Mast, M. S., Rammsayer, T. H., Hall, J. A., Murphy, N. A. (2020). A meta-analysis of the relationship between emotion recognition ability and intelligence. *Cogn Emot*, *34*, 329–351.

Simpson, G. L. (2023). *gratia: Graceful ggplot-based graphics and other functions for GAMs fitted using mgcv*. https://gavinsimpson.github.io/gratia/

Singh, G., Liskovoi, L., Launer, S., Russo, F. (2019). The emotional communication in hearing questionnaire (EMO-CHeQ): Development and evaluation. *Ear Hear*, *40*, 260–271.

Souza, P. (2016). Speech perception and hearing aids. In G. R. Popelka, B. C. J. Moore, R. R. Fay, A. N. Popper (Eds.) Hearing Aids, Springer Handbook of Auditory Research, vol 56. Springer International Publishing.

Souza, P. E., Arehart, K. H., Kates, J. M., Croghan, N. B. H., Gehani, N. (2013). Exploring the limits of frequency lowering. *J Speech Lang Hear Res*, *56*, 1349–1363.

Souza, P. E., Wright, R. A., Blackburn, M. C., Tatman, R., Gallun, F. J. (2015). Individual sensitivity to spectral and temporal cues in listeners with hearing impairment. *J Speech Lang Hear Res*, *58*, 520–534.

Steeneken, H. J. M., & Houtgast, T. (1980). A physical method for measuring speech-transmission quality. *J Acoust Soc Am*, *67*, 318–326.

Taumoepeau, M., & Ruffman, T. (2006). Mother and infant talk about mental states relates to desire language and emotion understanding. *Child Dev*, *77*, 465–481.

Tomblin, J. B., Oleson, J., Ambrose, S. E., Walker, E. A., McCreery, R. W., Moeller, M. P. (2020). Aided Hearing Moderates the Academic Outcomes of Children With Mild to Severe Hearing Loss. *Ear Hear*, *41*, 775–789.

Tomblin, J. B., Harrison, M., Ambrose, S. E., Walker, E. A., Oleson, J. J., Moeller, M. P. (2015). Language Outcomes in Young Children with Mild to Severe Hearing Loss. *Ear Hear*, *36*, 76S–91S.

van Rij, J., Wieling, M., Baayen, R. H., van Rijn, H. (2022). itsadug: Interpreting time series and autocorrelated data using GAMMs *(2.4.1) [Computer software]*.

Van Tasell, D. J. (1993). Hearing loss, speech, and hearing aids. J Speech Hear Res, *36*, 228–244.

Volkova, A., Trehub, S. E., Schellenberg, E. G., Papsin, B. C., Gordon, K. A. (2013). Children with bilateral cochlear implants identify emotion in speech and music. *Cochlear Implants Int*, *14*, 80–91.

von Eiff, C. I., Skuk, V. G., Zäske, R., Nussbaum, C., Frühholz, S., Feuer, U., Guntinas-Lichius, O., Schweinberger, S. R. (2022). Parameter-specific morphing reveals contributions of timbre to the perception of vocal emotions in cochlear implant users. *Ear Hear*, *43*, 1178–1188.

von Salisch, M., Haenel, M., Freund, P. A. (2013). Emotion understanding and cognitive abilities in young children. *Learn Individ Differ*, *26*, 15–19.

Way, I. F., Applegate, B., Cai, X., Franck, L. K., Black-Pond, C., Yelsma, P., Roberts, E., Hyter, Y., Muliett, M. (2010). Children's Alexithymia Measure (CAM): A new instrument for screening difficulties with emotional expression. *J Child Adolesc Trauma*, *3*, 303–318.

Wood, S. N. (2003). Thin-plate regression splines. *J R Stat Soc B*, *65*, 95–114.

Wood, S. N. (2004). Stable and efficient multiple smoothing parameter estimation for generalized additive models. *J Am Stat Assoc*, *99*, 673–686.

Wood, S. N. (2011). Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *J R Stat Soc B*, *73*, 3–36.

World Health Organization. (2021). *World Report on Hearing*. World Health Organization.

Wood, S. N. (2017). *Generalized Additive Models: An Introduction With R* (2nd ed.). Chapman and Hall/CRC.

Wood, S. N., Pya, N., Säfken, B. (2016). Smoothing parameter and model selection for general smooth models (with discussion). *J Am Stat Assoc*, *111*, 1548–1563.

Yeshoda, K., Raveendran, R., Konadath, S. (2020). Perception of vocal emotional prosody in children with hearing impairment. *Intern J Pedia*t *Otorhinolaryngol*, *137*, 110252.