



# Language-Independent Hearing Screening – Increasing the Feasibility of a Hearing Screening Self-Test at School-Entry

Trends in Hearing  
Volume 26: 1–15  
© The Author(s) 2022  
Article reuse guidelines:  
sagepub.com/journals-permissions  
DOI: 10.1177/  
23312165221122587  
journals.sagepub.com/home/tia  


Elien Van den Borre<sup>1</sup> , Sam Denys<sup>1</sup> , Lea Zupan<sup>2</sup>,  
Jan A. P. M. de laet<sup>3</sup>, Nina Božanić Urbančič<sup>4,5</sup>,  
Astrid van Wieringen<sup>1</sup> and Jan Wouters<sup>1</sup>

## Abstract

A tablet-based language-independent self-test involving the recognition of ecological sounds in background noise, the Sound Ear Check (SEC), was adapted to make it feasible for young children. Two experiments were conducted. The first experiment investigated the SEC's feasibility, as well as its sensitivity and specificity for detecting childhood hearing loss with a monaural adaptive test procedure. In the second experiment, the SEC sounds, noise, and test format were adapted based on the findings of the first experiment. The adaptations were combined with three test procedures, one similar to the one used in Experiment 1, one presenting the sounds dichotically in diotic noise, and one presenting all the sounds with a fixed signal-to-noise ratio and a stopping rule. Results in young children show high sensitivity and specificity to detect different grades of conductive and sensorineural hearing loss (70–90%). When using an adaptive, monaural procedure, the test duration was approximately 6 min, and 17% of the results obtained were unreliable. Adaptive staircase analyses showed that the unreliable results probably occur due to attention/motivation loss. The test duration could be reduced to 3–4 min with adapted test formats without decreasing the test-retest reliability. The unreliable test results could be reduced from 17% to as low as 5%. However, dichotic presentation requires longer training, reducing the dichotic test format's feasibility.

## Keywords

hearing screening, suprathreshold, childhood hearing loss, conductive hearing loss, sensorineural hearing loss

Received 22 February 2022; Revised 5 August 2022; accepted 11 August 2022

## Introduction

About one per 1000 children are born with hearing loss (Butcher et al., 2019). Untreated hearing loss increases the risk for speech, language, and learning difficulties and low social-communicative abilities (Bess et al., 1998; Porter et al., 2013; Winiger et al., 2016). Many high-income countries have implemented neonatal hearing screening (NHS) programs to ensure early detection and rehabilitation of congenital hearing loss, and a third of these countries cover at least 85% of newborns (Neumann et al., 2022).

However, reports indicate that the prevalence of hearing loss doubles before the age of nine (Fortnum et al., 2001). Late-onset, progressive, and acquired hearing loss are not detected with NHS; yet in most countries, systematic hearing screening does not exist beyond NHS (Fortnum et al., 2001; Mehra & Eavey, 2009). In countries and regions where school-age hearing screening (SHS) is implemented, testing methods (or procedures), referral strategies, and follow-up measures

can be different (Dettman et al., 2007; Moeller, 2000; Neumann et al., 2006; Yong et al., 2020; Yoshinaga-itano et al., 2001; Yoshinaga-Itano et al., 1998).

<sup>1</sup>Department of Neurosciences, Research Group ExpORL, KU Leuven, Leuven, Belgium

<sup>2</sup>Department of ear, nose, and throat, General Hospital Celje, Celje, Slovenia

<sup>3</sup>Department of Audiology (ear, nose, and throat), Leiden University Medical Center, The Netherlands

<sup>4</sup>Department of Otorhinolaryngology and Cervicofacial Surgery, University Medical Centre Ljubljana, Ljubljana, Slovenia

<sup>5</sup>Faculty of Medicine, Department of Otorhinolaryngology, University of Ljubljana, Ljubljana, Slovenia

## Corresponding author:

Elien Van den Borre, Department of Neurosciences, ExpORL, KU Leuven, Herestraat 49 bus 721, B-3000 Leuven, Belgium.  
Email: elien.vandenborre@kuleuven.be



A recent narrative review highlighted the importance of globally standardized school hearing screening protocols (Yong et al., 2020). The World Health Organization (WHO) and the European Federation of Audiology Sciences (EFAS) strongly recommend screening all children for hearing loss and ear diseases at school entry as a bare minimum (World Health Organization, 2021; Wouters et al., 2017). Ideally, every child with a pure-tone average (PTA) of 20 dB HL or higher should be identified, as even mild hearing loss can affect educational attainment (Moore et al., 2019). Identification of hearing loss allows for the provision of support and intervention, which has proven to have positive effects on the development of those children (Holzinger et al., 2022; Moeller, 2000; Pimperton et al., 2016; Watkin et al., 2007; Yoshinaga-Itano et al., 1998).

However, the few installed SHS programs at this moment are often not able to detect hearing loss lower than 30 – 35 dB, mainly because they rely on behavioral methods based on pure-tone audiometry. These methods only have reliable and valid outcomes under strictly controlled conditions. For example, pure-tone audiometric threshold detection is extremely sensitive to external noise. Even in the absence of clear environmental noise, ambient noise levels can easily be above 40 dBA, and performing reliable pure-tone audiometry is difficult when no soundproof booth or quiet room is present (Prieve et al., 2015). Unfortunately, these conditions are not always easy to meet, compromising the screening programs' quality.

A possible solution is speech-in-noise (SPIN) testing. SPIN tests present speech in background noise at a level well above the hearing threshold of normal-hearing (NH) participants. They aim to estimate a speech reception threshold (SRT), the signal-to-noise ratio (SNR) at which the listener correctly identifies a certain proportion of the target speech material. Poor speech understanding in noise is usually among the first complaints of hearing-impaired (HI) people (Kramer et al., 2008; Smith et al., 2005). Therefore, SPIN tests are highly sensitive to (mild) hearing losses and measure reliably even in the presence of some environmental noise (Plomp, 1985). Hearing loss can affect people's perceptions in two ways. Firstly, hearing loss can have an attenuation component, which causes soft sounds to be perceived more quietly, reducing their audibility. Secondly, a distortion component can be present even when it is loud enough to be heard. As SPIN tests are mostly suprathreshold, they measure the distortion component of hearing loss rather than the attenuation component, unlike pure-tone audiometric thresholds, which measure mostly the attenuation component. As a result, SPIN tests are most sensitive to sensorineural hearing loss (SNHL), where both an attenuation component and a distortion component are present, rather than to conductive hearing loss (CHL), where only an attenuation component is present (Plomp, 1978). However, when the noise level of an adaptive SPIN test is fixed, which is the case in many SPIN tests (Denys et al., 2019; Jansen et al., 2010;

Smits et al., 2004), the level of the stimuli can come close to the threshold level, increasing the impact of the attenuation component on the results.

A SPIN screening test that has become increasingly popular over the past fifteen years is the Digit Triplet Test (DTT) (for a review, see Van den Borre et al., 2021). The DTT was introduced by Smits et al. (2004) as an automated self-test targeting adults, in which combinations of three digits are presented in noise at adaptively varying levels. Research showed that the test has high test-retest reliability, a high correlation with average pure-tone thresholds and high sensitivity and specificity of approximately 80% to detect elevated pure-tone thresholds, which are essential characteristics of a high-quality screening test (Smits et al., 2004).

In Flanders, the Dutch-speaking part of Belgium, this test has been successfully implemented for large-scale systematic hearing screening of children from the 5<sup>th</sup> grade of elementary school (9 to 12 years old) and the 3<sup>rd</sup> grade of secondary school (13 to 16 years old) by school health services (Denys et al., 2018). However, the DTT is not ideal for an international school entry hearing screening (SEHS) standard. For children at school-entry (around six years old), SRTs obtained with a DTT-self-test compare poorly to SRTs obtained with an oral administration of the DTT, in which the children have to repeat the digits they hear, and a test administrator fills in the answers (Denys et al., 2021). Moreover, only a small proportion of young children obtain reliable SRTs (i.e., SRTs with a standard deviation, SD, comparable with the step size used during the adaptive test administration) and test durations are long. These issues show that the DTT is too challenging as a self-test at the age of school entry (Denys et al., 2021). On top of this, the DTT uses speech, namely digits, which are language-dependent. In Europe alone, 24 different official languages are spoken. This number increases dramatically when taking dialects into account. Before a DTT-version in a new language can be used, it needs to go through an extensive developmental and validation phase. Not all countries have the resources to develop this paradigm, making it difficult to use it as an across-country standard.

A possible solution that would overcome the issues experienced with the DTT without losing its advantages is an adaptive self-test that uses non-speech sounds with spectral characteristics similar to speech, represented by drawings or images. Non-speech sounds are language-independent and do not require the knowledge of numbers and their written forms. Additionally, using one sound instead of three digits presumably reduces the working memory load. The use of non-speech sounds to estimate hearing abilities has been researched before and was found to be a promising alternative to pure-tone audiometry both for adults (Myers et al., 1996) and children (Nolte et al., 2016).

In 2019, the Sound Ear Check (SEC), a sound-in-noise self-test, was described and optimized by Denys et al.

(2019). It is an automated adaptive closed-set tablet-based self-test based on the perception of ecological sounds (e.g., barking of a dog or honking of a car) presented in spectrally matched background noise. Moreover, the sounds cover the frequency range that is most important for speech understanding (500–4000 Hz; Fletcher & Galt, 1950; French & Steinberg, 1947) and identification of the sounds of the SEC is more likely to be reliant on the temporal properties of the sounds than, for example, the digits used in the DTT (Denys et al., 2019). As the temporal processing may be distorted when CHL is present (Balen et al., 2009; Bayat et al., 2017; Hartley & Moore, 2003; Moore et al., 2003), the identification of the sounds of the SEC could be affected by CHL as well. This may give the test sensitivity and specificity to CHL, on top of the sensitivity and specificity to SNHL.

Sensitivity and specificity rates of approximately 80% were determined to detect adults with slight SPIN recognition difficulties. Additionally, a high test-retest reliability of 1 dB was obtained. Piloting the test in young NH children (5–7 years) showed test-retest reliability of 1.3 dB, which was promising.

However, the feasibility of SEC described by Denys et al. (2019) for large-scale SEHS is questionable. First, the test duration is 5–6 min. In large-scale screening protocols for children, multiple classes often need to be tested in one day, sometimes combined with other medical examinations at the same moment. Implementing a hearing screening that takes 5 to 6 min in such protocols can be problematic. Additionally, about 20% of the children could not complete the test reliably, possibly due to an overly long test duration, a lack of engagement, or the use of eight sounds which could be too taxing for children. Taken together, the concept of the SEC is very promising, but the test needs optimization for SEHS internationally. Additionally, the SEC's sensitivity and specificity to detect childhood hearing loss need to be investigated.

In this article, two different experiments are reported. In the first experiment, three research questions are addressed: (1) What is the sensitivity and specificity of the SEC as described by Denys et al. (2019) to detect different grades of hearing loss in young children?; (2) Is the SEC feasible as a SEHS in different countries regarding the duration, the familiarity of the non-speech sounds, and the number of unreliable results?; (3) What aspects of the SEC limit its feasibility as a SEHS?

The SEC was then adapted based on the findings of Experiment 1, and the new test procedures were evaluated in a different population of children in Experiment 2, which addresses three research questions: (1) Do the new, adapted test procedures reduce the test duration of the SEC? (2) Do the new, adapted test procedures show comparable test-retest reliability to the reference test procedure? (3) Do the adaptations of the test procedures solve the limitations of the original SEC procedure?

With this research, we aim to take another step towards developing an internationally applicable self-test for SEHS.

## Experiment 1

### Participants

NH and HI children were recruited. All children attended the 1<sup>st</sup> and 2<sup>nd</sup> grades of elementary school (5 years 8 months – 7 years 11 months), and were recruited in Flanders, the Netherlands, and Slovenia. Ears were considered separately, meaning they could be included in different hearing loss groups in case of asymmetric hearing loss or an unreliable result in one of the ears. Results were included in the NH group when that ear had an air conduction (AC) PTA<sub>0.5,1,2,4kHz</sub> < 20 dB HL. Results were included in the CHL group when an AC PTA<sub>0.5,1,2,4kHz</sub> ≥ 20 dB HL, a bone conduction (BC) PTA<sub>0.5,1,2,4kHz</sub> < 20 dB HL, and an Air Bone Gap (ABG) ≥ 10 dB were present. Results were included in the SNHL group, when an AC PTA<sub>0.5,1,2,4kHz</sub> ≥ 20 dB HL, a BC PTA<sub>0.5,1,2,4kHz</sub> ≥ 20 dB HL, and an ABG < 10 dB were present. Results were included in the mixed hearing loss (MHL) group when an AC PTA<sub>0.5,1,2,4kHz</sub> ≥ 20 dB HL, a BC PTA<sub>0.5,1,2,4kHz</sub> ≥ 20 dB HL, and an ABG > 10 dB were present.

In total, 215 children were recruited and did the full protocol. The results of three ears measured in three children were excluded because hearing loss was present, and no bone conduction thresholds were measured. The results of one ear were removed because no reliable audiogram could be obtained in that ear. Unreliable SRT results, i.e., SD of the SRT > 3 dB, were measured in 71 ears. These results were included in Group 1-unreliable. Group 1-reliable consisted of results of 355 ears. The number of ears included in Group 1-reliable and Group 1-unreliable and the average age per hearing group and country are given in Table 1. The mean PTA<sub>0.5,1,2,4kHz</sub> for the NH, SNHL, MHL and CHL group were 10 ± 6 dB HL, 40 ± 13 dB HL, 44 ± 14 dB HL and 32 ± 11 dB HL, respectively. The average audiogram and SD for the different hearing groups is given in Figure 1.

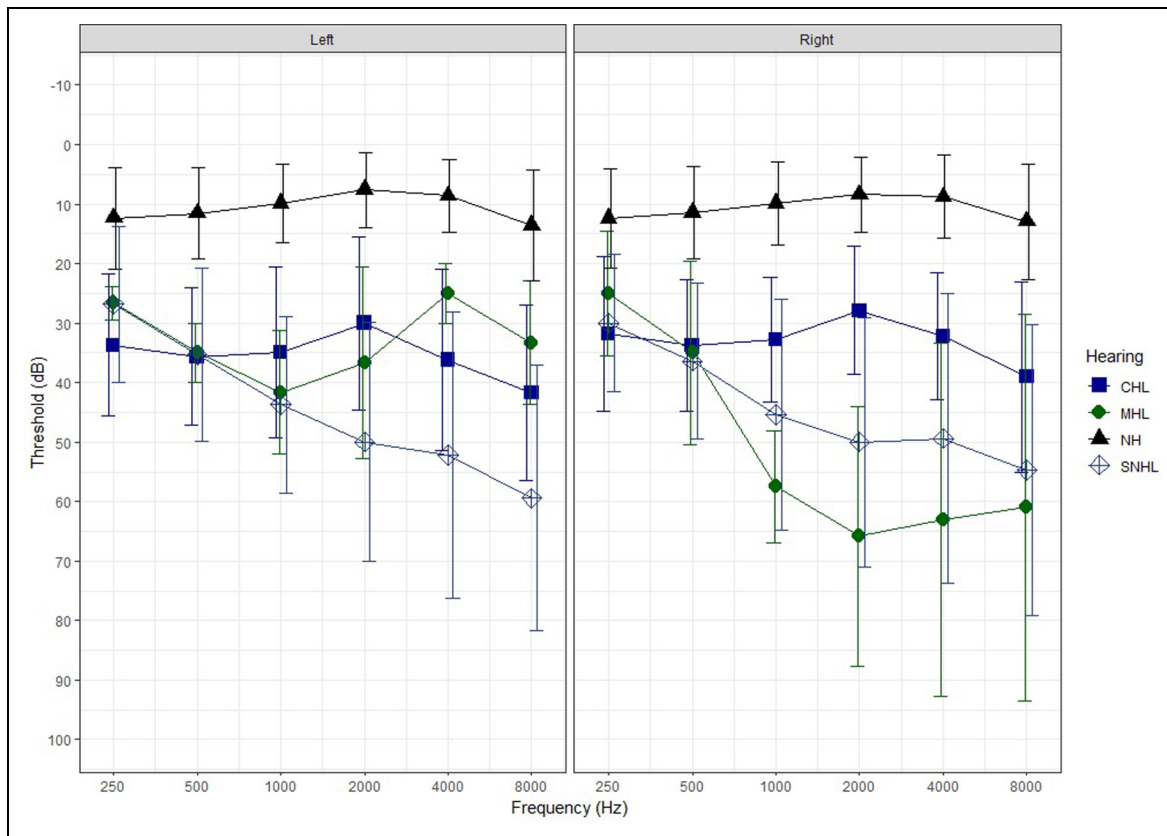
## Methods

### Materials

In Flanders, a Madsen Orbiter 922 or Madsen Midimate 622 audiometer connected to an HDA-200 headphone, calibrated according to ISO-389 standards, was used. Partners in Slovenia used an Interacoustics Clinical Audiometer AC40. Partners in The Netherlands used an Interacoustics Equinox 2.0 audiometer and an Interacoustics Callisto portable audiometer. The SEC was administered on a 7-inch Samsung Galaxy Tab A tablet, connected to DD65 transducers embedded in Peltor caps. The test setup was calibrated to the sound-weighted

**Table 1.** Number of Ears and the Average age per Group in Experiment 1 for The Netherlands (NL), Belgium/Flanders (BE (FL)) and Slovenia (SL).

Experiment 1	Group 1 (Ears)					
	Reliable			Unreliable		
Country	NL	BE (FL)	SL	NL	BE (FL)	SL
NH	26 (77 ± 7 months)	101 (82 ± 8 months)	112 (81 ± 8 months)	6 (74 ± 7 months)	19 (78 ± 6 months)	20 (78 ± 6 months)
CHL	9 (85 ± 6 months)	16 (79 ± 8 months)	45 (79 ± 7 months)	0	3 (84 ± 10 months)	11 (75 ± 3 months)
SNHL	4 (85 ± 5 months)	19 (80 ± 10 months)	16 (83 ± 8 months)	1 (91 months)	4 (73 ± 9 months)	5 (86 ± 4 months)
MHL	3 (77 ± 4 months)	0	4 (78 ± 2 months)	2 (84 ± 5 months)	0	0

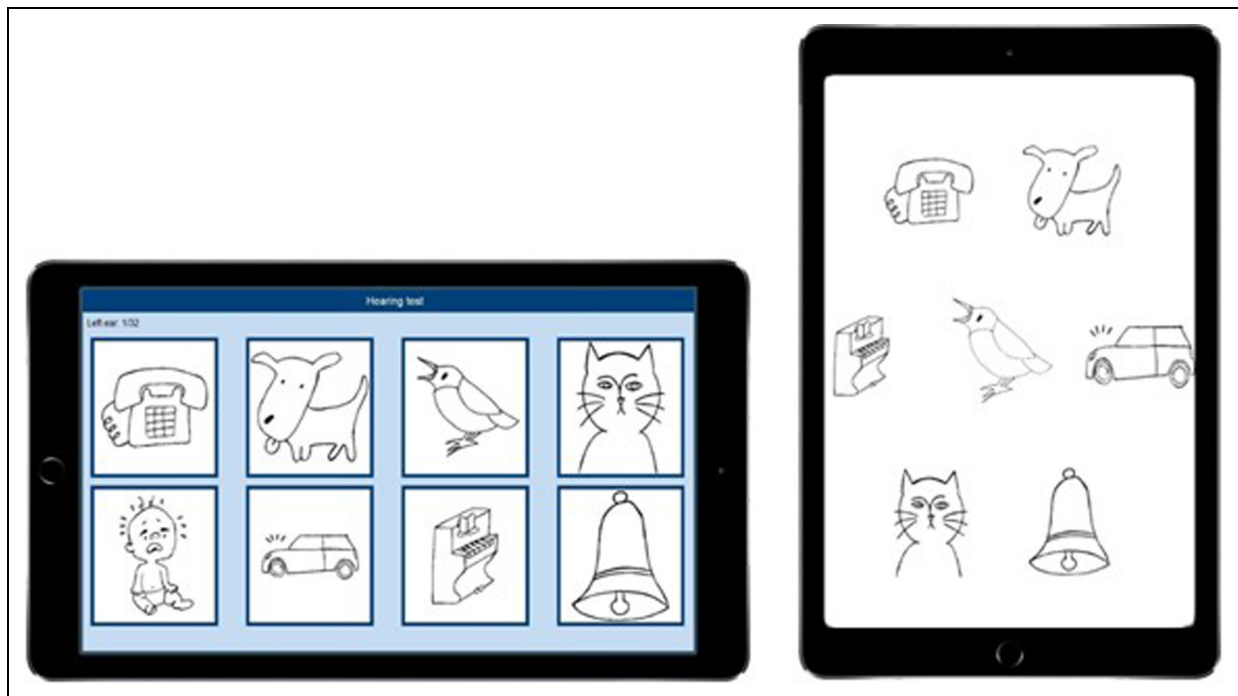
**Figure 1.** Average audiogram and SD for the different hearing groups in experiment 1.

noise (SWN), i.e., noise with a spectrum similar to the average spectrum of the test materials, at 80 dB Sound Pressure Level (SPL) with a Brüel & Kjaer Sound level meter 2260 and a Brüel & Kjaer 4153 artificial ear using the flat plate. This study was approved by the Ethics Committee Research UZ / KU Leuven. The SEC test used and the protocol for children are described in the following paragraphs.

### Sound Ear Check Version

The original procedure (SEC) was based on the procedure described by Denys et al. (2019). In short, eight sounds, e.g., the barking of a dog or the honking of a car, were presented in SWN with a fixed level of 65 dB SPL. Eight drawings, chosen based on recognizability (Denys et al., 2019),

were shown on the tablet screen (Figure 2). The SEC consisted of three subsequent phases. The first phase was the acclimatization phase. In this phase, the sounds were presented diotically at a fixed SNR of 10 dB SNR. Each sound was presented in random order until identified correctly, after which the image disappeared from the screen as a sign that the sound-image mapping was correct. The second phase was a diotic training phase consisting of 24 trials with every sound played three times ( $3 \times 8$  sounds) in random order. From this phase on, the images did not disappear after a correct answer, so the participant did not get direct feedback. An adaptive procedure was used, starting at an SNR of 0 dB. The maximal SNR was 10 dB SNR, meaning that when the adaptive procedure reached 10 dB SNR and was followed by another incorrect answer, the



**Figure 2.** Interface of the original SEC (left) used in experiment 1 and SEC<sub>REF</sub>, SEC<sub>APH</sub>, and SEC<sub>FIX</sub> (right) used in experiment 2.

SNR did not change. The stimuli were adapted in steps of 2 dB. For the first seven trials, a 1-up-1-down procedure was used to converge fast to the SRT. A 1-up-2-down procedure was used for the subsequent trials, targeting a recognition probability of 71% correct (Levitt, 1971). The third phase was the actual test phase, which used 24 trials and the same procedure as the training phase but with monaural stimulus presentation, first testing the left ear, then the right ear. The SRT was calculated by averaging the SNR of the last 17 trials, including a non-presented (imaginary) 25th trial of which the SNR was calculated based on the identification response of the final presented item.

### Protocol

Children were tested at home, in a hospital, or at school. In all settings, all measurements were conducted in a quiet room where visual and auditory distractions could be kept minimal. Moreover, the HDA-200 headphones that were used in the experiments provided shielding from ambient noise. The protocol started with audiometry. The hearing thresholds (250–8000 Hz) were measured according to the Hughson-Westlake method. If AC thresholds (250–8000 Hz) were > 20 dB HL, BC thresholds were measured (250–4000 Hz). Next, the SEC was performed. The test was performed as a self-test without extra help from the researcher. No additional instructions were given during the test, except when the child asked a question or stopped continuing. All children in Group 1 were asked their

opinion about the test. Moreover, test leaders documented aspects that reduced the autonomy of the children during the self-test.

### Statistical Analyses

Statistical analyses were performed using R and R studio (RStudio, 2020). Normality was checked with Shapiro-Wilk normality tests. Differences between data collected in different countries on the SRT-values were investigated with a one-way analysis of variance (ANOVA) test on the results of the NH group to ensure no differences were present. SRT-values in Experiment 1 were normally distributed. A possible age effect was investigated with a simple linear regression model on the SRT-values of the NH ears. Only reliable results were used to investigate the age effect and the effect of country.

To investigate the first research question, namely “*What is the sensitivity and specificity of the SEC as described by Denys et al. (2019) to detect different grades of hearing loss in young children?*” receiver operating characteristic (ROC) analyses were conducted based on the data of Group 1-reliable. Sensitivity and specificity to detect hearing losses with a  $PTA_{0.5,1,2,4kHz} > 20$  dB HL,  $> 30$  dB HL, and  $> 40$  dB HL were estimated. ROC analyses were constructed for CHL and SNHL separately. As only a limited number of children with MHL were tested, no independent analyses were done for MHL. One ROC analysis was done for CHL, MHL and SNHL combined. The

assumptions of ROC analyses, namely that the measurement of interest is continuous with an independent diagnosis, that the state variable is independent of the measurement of interest and that the cases are a random sample, were all met.

To investigate the second research question, namely “*Is the SEC feasible as a SEHS in different countries regarding the duration, the familiarity of the non-speech sounds, and the number of unreliable results?*” three different aspects were analyzed. Firstly, the duration of the SEC and its SD were estimated. Secondly, a confusion matrix was made based on the NH ears in Group 1-reliable data to determine the familiarity of the drawings or sounds in different countries. Lastly, the number of unreliable results was estimated. Every SRT-value with an SD above 3 dB was classified as unreliable.

Two different aspects were analyzed to investigate the third research question: “*What aspects of the SEC limit its feasibility as a SEHS?*”

Firstly, the adaptive staircases of the NH ears of Group 1-unreliable were investigated and compared to the adaptive staircases of the NH ears of Group 1-reliable to determine why they became unreliable. Both the test flow of the results and the response times were compared between Group 1-reliable and Group 1-unreliable. The test flow of the unreliable results was determined by dividing the adaptive staircases into two parts (trial 7 to 16 and trial 16 to 24). Based on the difference between the SRTs calculated from these two parts, staircases were divided into three groups; namely an increasing pattern (the SRT of the second part is > measurement error (1.3 dB) higher than the SRT of the first part), a downward sloping pattern (the SRT of the second part is > measurement error (1.3 dB) lower than the SRT of the first part), and a horizontal pattern (the SRT of the first part and the second part differ less than the measurement error but there is general instability of the staircase, meaning that no specific increasing or decreasing trend is visible, yet the SD of the SRT is above 3 dB). The response times were compared with an ANOVA. The differences and the response times were normally distributed.

The second analysis to answer the third research question was a qualitative analysis of aspects that negatively affect the autonomy of the children during testing. Researchers were asked to document these aspects, and the children’s opinion was asked and documented, so a detailed inventory of all encountered problems could be made.

## Results

The average SRT-value for NH children was  $-8.7 \pm 2.1$  dB SNR. The average SRT-values for SNHL, CHL and MHL were respectively  $-4.3 \pm 4.6$  dB SNR,  $-7.2 \pm 3.2$  dB SNR, and  $-3.3 \pm 8.0$  dB SNR. The SRT-values per hearing group are visualized in Figure 3. No differences were found between data collected in different countries

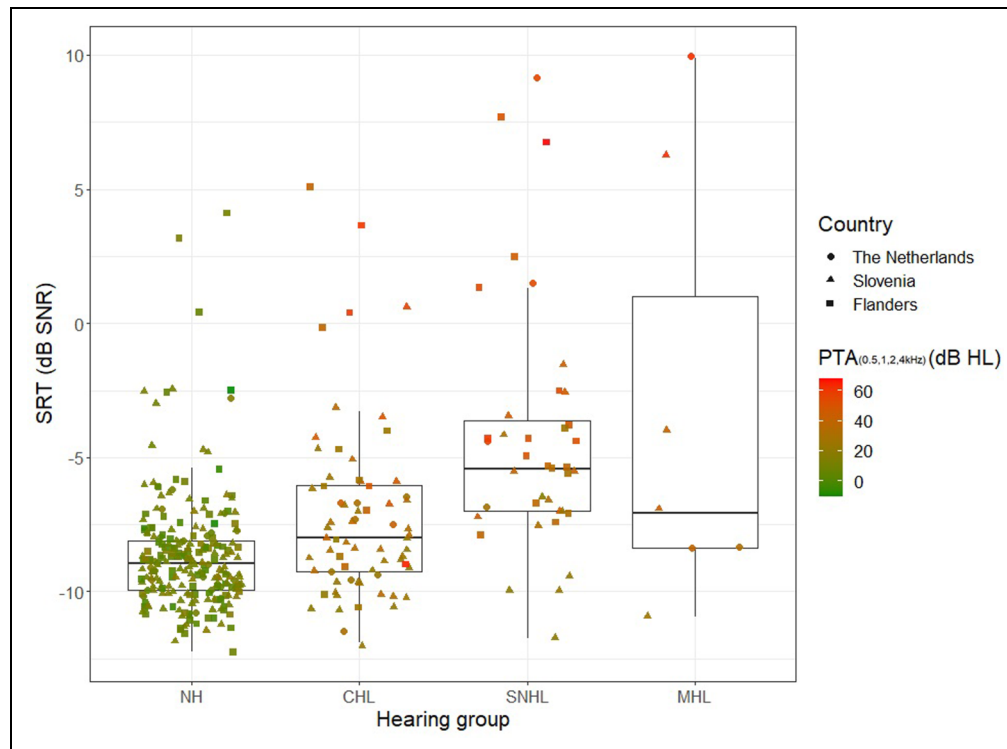
( $F(2, 236) = 0.6, p = 0.5$ ). Age had no significant effect on the SRT-values ( $t(237) = -1.5, p = 0.1$ ). Firstly, the SEC had very good sensitivity and specificity to detect SNHL > 20 dB (0.80–0.90). For CHL, very good sensitivity to pick up CHL > 40 dB (0.80–0.90) and reasonable sensitivity to identify CHL > 20 dB (0.60–0.70) were determined. To detect SNHL, CHL and MHL together, reasonable sensitivity and specificity to detect hearing loss > 20 dB HL (0.60–0.70) and very good sensitivity and specificity to detect hearing loss > 40 dB HL (0.80–0.90) were determined. The area under the curve (AUC), optimal pass-fail criteria and the sensitivity and specificity are given in Table 2.

Secondly, the average test duration for the NH ears in Group 1-reliable, including acclimatization, training, the left, and the right test, was 5 min 50 s  $\pm$  1 min 14 s. Figure 4 shows that children could identify most sounds accurately, as indicated by the high diagonal scores and confusions below the chance level (12.5%), except for the baby, which was often confused with the cat (20%). This confusion was unidirectional; the cat-baby confusion occurred only in 5% of presentations, which was well below the chance level. This confusion was present in both Slovenia and The Netherlands/Flanders. Children reported hearing an “angry cat” instead of a “crying baby.”

Documented limitations indicated that children sometimes reported hearing some of the sounds, e.g., the piano, inside the noise sound. Moreover, some children seemed to experience difficulties with guessing and therefore waited long or until they received new instructions before answering. This resulted in inflated response times. The average response time for NH children was 4.4 s, and the SD was 3.1 s. 22 NH children had response times > 9.5 s on one or more trials, which was more than 1.65 SD above the average.

Lastly, approximately one out of six of the results are unreliable. The average response time for the NH children in Group 1-reliable was not significantly different from the NH children in Group 1-unreliable ( $F(1, 6814) = 0.2, p = 0.62$ ). The average response times were respectively  $4.4 \pm 3.1$  s and  $4.3 \pm 3.3$  s for the NH in Group 1-reliable and Group 1-unreliable.

Seven out of 45 unreliable tests showed a decreasing trend of the adaptive staircase (4 left, 3 right). Twenty-six out of 45 unreliable results (12 left, 14 right) showed an increasing trend. Twelve results showed general instability of the staircase (3 left, 9 right). The unreliable results, divided into the different possible trends, are shown in Figure 5. The average adaptive staircase showed an increase in SNR from trial 8 on for Group 1-unreliable. For Group 1-reliable, the difference between the SRT calculated based on trials 7 to 16 and the SRT calculated based on trials 16 to 24 was  $0.3 \pm 2.0$  dB SNR. For Group 1-unreliable, this was  $3.6 \pm 5.6$  dB SNR. This was a significant difference ( $F(1, 282) = 49, p < 0.001$ ). The average staircases are shown in Figure 6.



**Figure 3.** Boxplot of SEC SRTs for different hearing groups. The shape of the dots visualizes the country, the color of the dots visualizes the PTA (0.5, 1, 2, 4 kHz).

### Optimization of the SEC

This section describes several modifications that were made to the original SEC to deal with some of the problems identified in Experiment 1. The sound of the baby appeared to be more complex than the other sounds used and was often confused with the sound of the cat. Where the baby was confused with the cat in 20% of the cases, on average, the other sounds were only confused 3% of the time. Earlier research already pointed to difficulties in identifying the baby sound (Denys et al., 2019). On top of this, children reported hearing some of the SEC stimuli in the noise, and some children showed problems with guessing when they were unsure of what they heard.

One out of six children obtained an unreliable result when performing the test without the researcher's help, making it challenging to do as a self-test. The adaptive staircases of these children showed that the majority could converge to low SNRs but performed worse towards the end of the test. This trend indicates attention or motivation loss.

The baby sound was removed to optimize the procedures, resulting in a test with seven different sounds and drawings. The interface of the adapted SEC versions is shown in Figure 2. New SWN was generated based on the seven included sounds and was then smoothed with finite impulse response filters to eliminate the recognizability of certain sounds in the noise (Henry et al., 2005). The spectra from

the unsmoothed SWN and the smoothed SWN are shown in Figure 7. A third-octave spectrum was made of the three noises; the original noise based on eight sounds used in the SEC, the unsmoothed noise based on seven sounds, and the smoothed noise based on seven sounds. A third-octave spectrum represents the critical bands of the human hearing (Danahauer et al., 1977). The mean absolute difference across bands (center frequencies 500–4000 Hz) between three noises was only  $1.1 \pm 0.9$  dB. Therefore, the masking capacity was expected to differ minimally (Greenwood, 1961).

The adapted test procedures proceeded automatically to avoid children being stuck when they did not want to guess. The time the test took to proceed was 1.65 SDs above the average answer time, estimated in Experiment 1, rounded to a full second (10 s). In the adapted versions, all the drawings moved during the presentation of the sounds to draw the child's attention to the screen while the sounds were played. All these optimizations were implemented in all the used tests in Experiment 2.

Two new procedures were implemented next to the original monaural adaptive procedure, namely a fixed SNR procedure (SEC<sub>FIX</sub>) and an antiphase procedure (SEC<sub>APH</sub>). The SEC<sub>FIX</sub> was based on a procedure suggested by Smits (Smits, 2017). This test procedure used a fixed SNR, i.e., the same SNR for all stimulus presentations and a variable number of trials. The stopping rule to calculate the number of trials,

**Table 2.** Values for Sensitivity and Specificity for CHL, SNHL and Combined CHL, SNHL and MHL.

PTA <sub>0.5,1,2,4</sub> kHz (dB HL)	CHL			
	AUC (95% CI)	Cut-off (dB SNR)	Sens	Spec
20	0.66 (0.59–0.72)	-8.6	0.61	0.62
30	0.77 (0.69–0.84)	-8.2	0.74	0.72
40	0.89 (0.83–0.94)	-6.8	0.80	0.86
PTA <sub>0.5,1,2,4</sub> kHz (dB HL)	SNHL			
	AUC (95% CI)	Cut-off (dB SNR)	Sens	Spec
20	0.86 (0.79–0.92)	-7.4	0.86	0.83
30	0.93 (0.91–0.96)	-7.4	0.97	0.81
40	0.95 (0.92–0.97)	-7.1	0.95	0.82
PTA <sub>0.5,1,2,4</sub> kHz (dB HL)	CHL + SNHL + MHL			
	AUC (95% CI)	Cut-off (dB SNR)	Sens	Spec
20	0.73 (0.68–0.78)	-8.4	0.69	0.65
30	0.84 (0.80–0.89)	-7.7	0.79	0.78
40	0.92 (0.89–0.95)	-7.1	0.90	0.81

is further described in the method section of Experiment 2. This procedure aims to substantially reduce the duration of the test, reducing the risk of attention and motivation loss, potentially resulting in fewer unreliable test results and easier practical implementation.

The SEC<sub>APH</sub> presented antiphasic stimuli in phasic noise. This presentation mode relied on binaural unmasking or Binaural Masking Level Differences (BMLD)(George et al., 2012). BMLDs are an improvement of the SRT in noise due to a phase difference of the stimuli between both ears. BMLDs are reported to be poorer for listeners with various SNHL and CHL types than NH controls. Moreover, BMLDs are reported to be poorer for listeners with asymmetrical hearing loss (de Sousa et al., 2020). As a result, asymmetrical hearing loss can be detected despite testing binaurally. de Sousa et al. (2020) showed that using antiphasic stimuli results in SRTs that are more strongly correlated with the worse ear PTA than with the better ear PTA and the sensitivity and specificity rates increase compared to a diotic test procedure. As for SEC<sub>FIX</sub>, testing binaurally was presumed to result in a shorter test. The new antiphasic condition was piloted on a group of 17 NH adults, who also completed a monaural version of the test. The aim of this pilot testing was to confirm that SRTs were better in the antiphasic condition, that testing time was shorter, that test-retest reliability was similar and the new test platform with moving drawings was not distracting or unclear.

## Experiment 2

### Participants

Children from the 1<sup>st</sup>, 2<sup>nd</sup>, and 3<sup>rd</sup> grades of elementary school (6 years – 9 years 4 months) were recruited in Flanders. Children with known hearing loss were excluded. In this experiment, 83 children were tested. One child was excluded because it was impossible to obtain a reliable audiogram. The mean PTA<sub>0.5,1,2,4kHz</sub> was 6 ± 6 dB HL. Five children obtained at least one unreliable SRT.

## Methods

### Materials

The same material as described for Flanders in Experiment 1 was used.

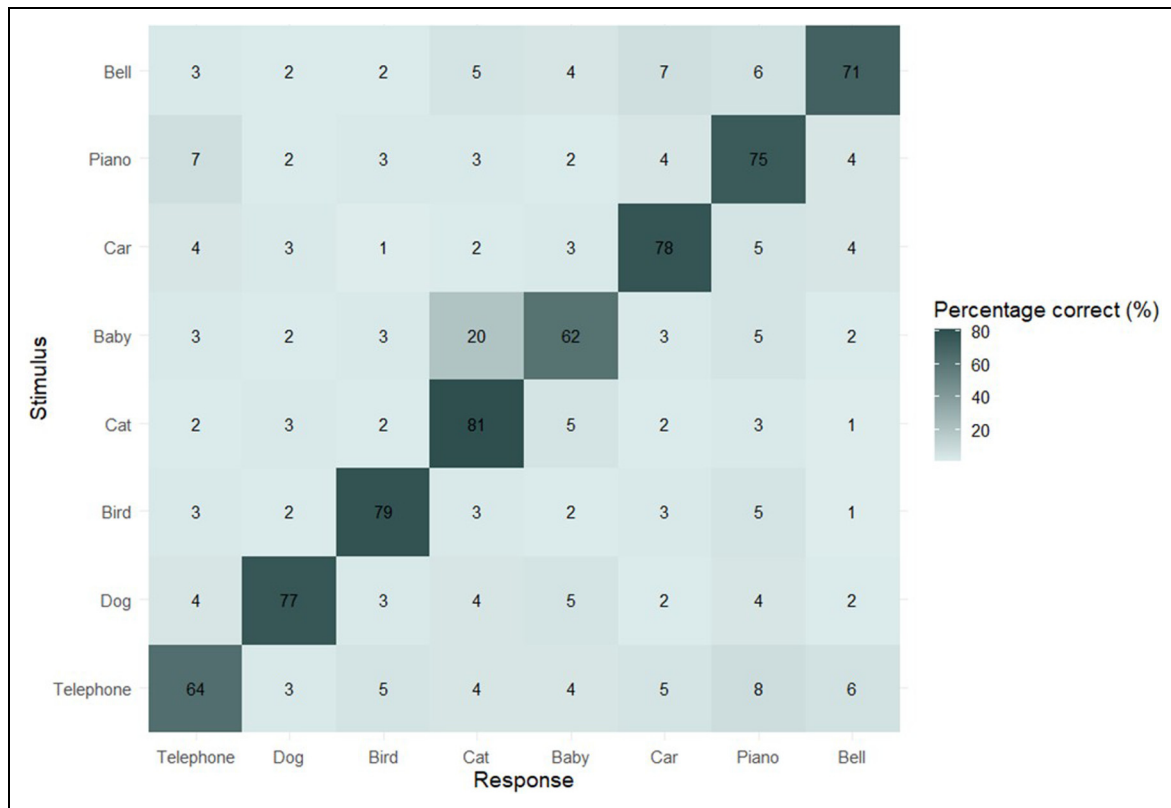
### Sound ear Check Versions

*An adapted reference procedure.* The adapted reference procedure (SEC<sub>REF</sub>) used the same overall test procedure as the SEC described in Experiment 1, but combined with new, adapted aspects. The changes are described and explained in the paragraph “Optimization of the SEC.”

*A fixed SNR procedure with a stopping rule.* The SEC<sub>FIX</sub> consisted of a 7-trial acclimatization phase, a training phase with a variable number of trials and two monaural test phases with a variable number of trials. The stopping rule used for the SEC<sub>FIX</sub> is based on the number of correct and false answers given (Smits, 2017). For each trial, a preliminary proportion correct was calculated. This is the number of correct answers divided by the total number of run-through trials. The proportion correct was compared with a predefined proportion. Then the probability of a pass and a refer result were calculated. 0.71 was used as the targeted proportion correct for a 1-up-2-down procedure. The test stopped as soon as one of those probabilities was higher than a predefined value, for example, 90%, or when the maximal number of trials was reached (N = 21 in this case). The probability distribution of p was estimated using Bayes’ theorem (Koch, 2007) to estimate the cumulative probabilities. The estimated probability distribution of p, f(p|n, k) depends on the number of correct responses k, the total number of presentations n, and the prior distribution. The cumulative probability that p > 0.71 is  $P(p > 0.71) = \int f(p|n, k)$ . Then the posterior distribution is given by the formula  $f(p|n, k) = \frac{(n+1)!}{k!(n-k)!} p^k (1-p)^{(n-k)}$  and  $P(p < 0.7) = 1 - P(p > 0.7)$ . The result was no longer an SRT, but a pass/fail classification.

The SNR used during the training phase was -7.1 dB SNR, which is the optimal cut-off to differentiate between NH children and children with CHL/SNHL > 40 dB. The training phase used the stopping rule described above but was slightly adapted. The training phase could only result in an early fail result and NH children who performed well would always complete the entire training needed. There are two reasons for this adaptation. Children who obtained a fail result in the training phase were unlikely to succeed in the test phase as the SNR used in the test phase was -7.8 dB SNR. This SNR is the optimal cut-off to differentiate between NH children and children with CHL > 30 dB and SNHL > 20 dB. It is 0.7 dB SNR lower compared to the SNR used during the training phase. Letting children complete a full training phase at an SNR firmly under their SRT would have been very demotivating. However, for children with SRT-values close to the SNR





**Figure 4.** Confusion matrix for sound ear check sounds calculated from results obtained in NH ears in three countries. Results from different countries are visualized together as no significant differences were present between countries.

used during the test phase, a full training phase was needed as the pass-fail criteria were estimated based on children who were already trained. If the children went into the test phase with less training, the pass-fail criteria were likely too challenging, resulting in more false-positive results.

After the training phase, a monaural test phase was performed for each ear with the stopping rule described above. The same optimizations were applied as for the  $SEC_{REF}$ .

**An antiphasic, adaptive procedure.** The  $SEC_{APH}$  started with the same acclimatization phase as described for the other  $SEC$  versions. It was followed by a dichotic training phase with 21 trials and a dichotic test phase with 21 trials, resulting in a total of 49 trials. Both the training and the test phase used the same adaptive procedure as described for the  $SEC_{REF}$  but with a starting SNR of  $-10$  dB SNR and noise fixed at 70 dB SPL because an antiphasic advantage of  $\pm 10$  dB SNR was expected (de Sousa et al., 2020; Smits et al., 2016). The same optimizations were applied as for the  $SEC_{REF}$ .

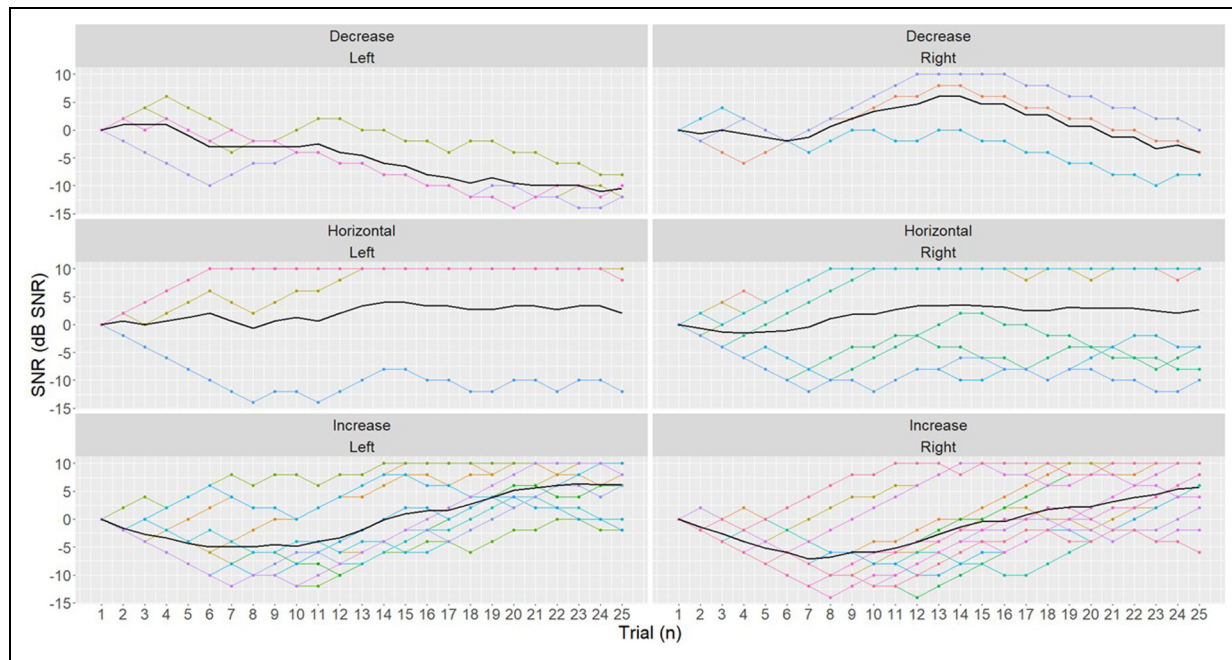
### Protocol

Children participating in Experiment 2 were tested at home in a quiet room. All children started with audiometry. The

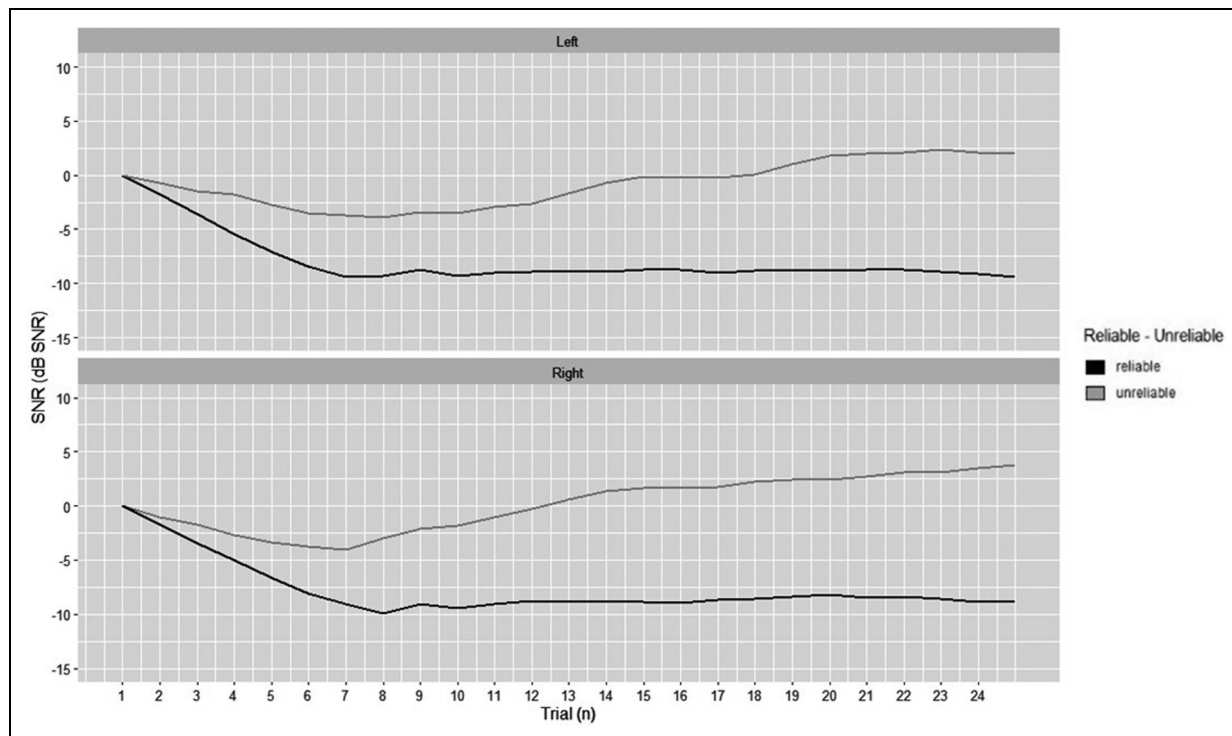
hearing thresholds (250–8000 Hz) were measured according to the Hughson-Westlake method. After audiometry, they did a test and retest run of the  $SEC_{REF}$  and two runs of either the  $SEC_{FIX}$  or the  $SEC_{APH}$ . Twenty-four children started with the new procedures ( $SEC_{FIX}/SEC_{APH}$ ). Fifty-eight children started with the  $SEC_{REF}$ . The instruction was the same as during the first experiment, and the same level of support was given. The instruction was not repeated between the first and the second test.

### Statistical Analyses

To investigate the first question of Experiment 2, namely “*Do the new, adapted test procedures reduce the test duration of the  $SEC$ ?*”, differences in duration were investigated with nonparametric tests as the duration data of the  $SEC_{REF}$  were not normally distributed. A Friedman test was used combined with a Wilcoxon rank-sum test. Moreover, to estimate whether one training phase is sufficient or more training is needed, which would increase the total test duration, a Friedman test was conducted on the first and the second test of both the  $SEC_{REF}$  and the  $SEC_{APH}$ . The Wilcoxon rank-sum test was used for pairwise comparisons. A non-parametric test was used because the SRT of the  $SEC_{APH}$  dichotic test 2,  $SEC_{REF}$  left 2 and  $SEC_{REF}$  right 2 were not normally distributed.



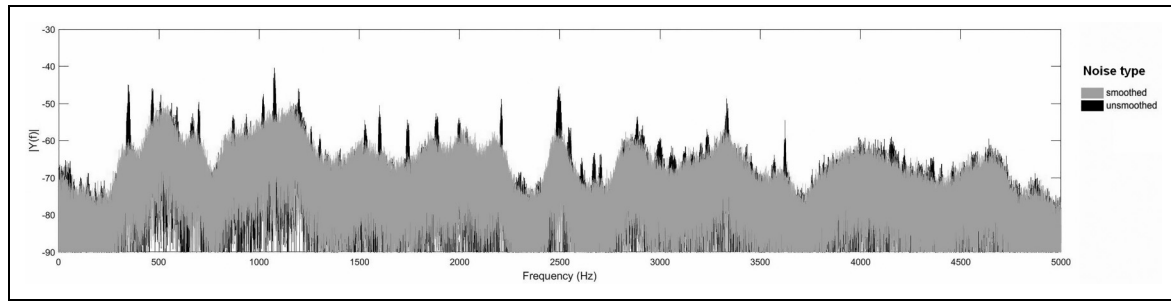
**Figure 5.** Trends of unreliable adaptive staircases in NH children: decrease (n = 7), horizontal (n = 12) and increase (n = 26). The bold line visualizes the average staircase within that specific group.



**Figure 6.** Average staircase of the NH ears in group I-unreliable and group I-reliable test results.

To investigate the second research question of Experiment 2, namely “Do the new, adapted test procedures show comparable test-retest reliability to the reference test procedure?” the

measurement error was estimated using the formula  $\frac{1}{\sqrt{2}} \sqrt{\frac{\sum (\Delta_k - \bar{\Delta}_k)^2}{n}}$ . In this formula, the SD of the differences



**Figure 7.** Smoothed and unsmoothed noise generated based on seven sounds visualized in a spectrum.

between test and retest is calculated and divided by  $\sqrt{2}$ .  $K$  stands for the differences between test and retest, and  $n$  is the number of tests done (Smits & Houtgast, 2005). This formula balances out the learning effect.

To investigate the last research question of Experiment 2, *“Do the adaptations of the test procedures solve the limitations of the original SEC procedure?”*, firstly, the number of unreliable test results was investigated. Test results for the adaptive procedures ( $SEC_{REF}$  and  $SEC_{APH}$ ) were classified as unreliable when the SD was larger than 3 dB. To calculate the percentage of unreliable test results, only the unreliable tests of the first test and not of the retest were counted to avoid counting unreliable test results due to fatigue because of the long protocol. Moreover, this allowed for a better comparison with the results of Experiment 1. Secondly, a qualitative analysis comparable to the analysis done in Experiment 1 was performed on aspects that could limit the feasibility of the SEC tests for implementation as a SEHS.

## Results

The average SRT for the  $SEC_{REF}$  was  $-8.7 \pm 1.6$  dB SNR, and the average SRT for the  $SEC_{APH}$  was  $-15.5 \pm 2.4$  dB SNR. The average duration of the reference procedure was  $4 \text{ min } 35 \text{ s} \pm 52 \text{ s}$ . This duration included seven trials in the acclimatization phase, 21 trials in the training phase, and  $2 \times 21$  trials test phase. The average duration of the  $SEC_{APH}$  and the  $SEC_{FIX}$  were respectively  $3 \text{ min } \pm 43 \text{ s}$  and  $3 \text{ min } 51 \text{ s} \pm 1 \text{ min } 5 \text{ s}$ . For the  $SEC_{APH}$  this included seven trials in the acclimatization phase, 21 trials in the training phase, and 21 trials in the test phase. The average number of trials in the fixed procedure was  $48 \pm 5$  trials. This included seven trials in the acclimatization phase and a variable number of trials in the training phase and the two test phases. The duration of the  $SEC_{APH}$  and the  $SEC_{FIX}$  did not differ significantly ( $Z = -2.01$ ,  $p = 0.05$ ). The duration of the  $SEC_{APH}$  and the  $SEC_{REF}$  differed significantly ( $Z = -4.35$ ,  $p < 0.01$ ) and the duration of the  $SEC_{FIX}$  and the  $SEC_{REF}$  differed significantly ( $Z = -1.98$ ,  $p = 0.04$ ). In addition, a significant difference was determined between the  $SEC_{REF}$  and the SEC ( $Z = -8.7$ ,  $p < 0.01$ ).

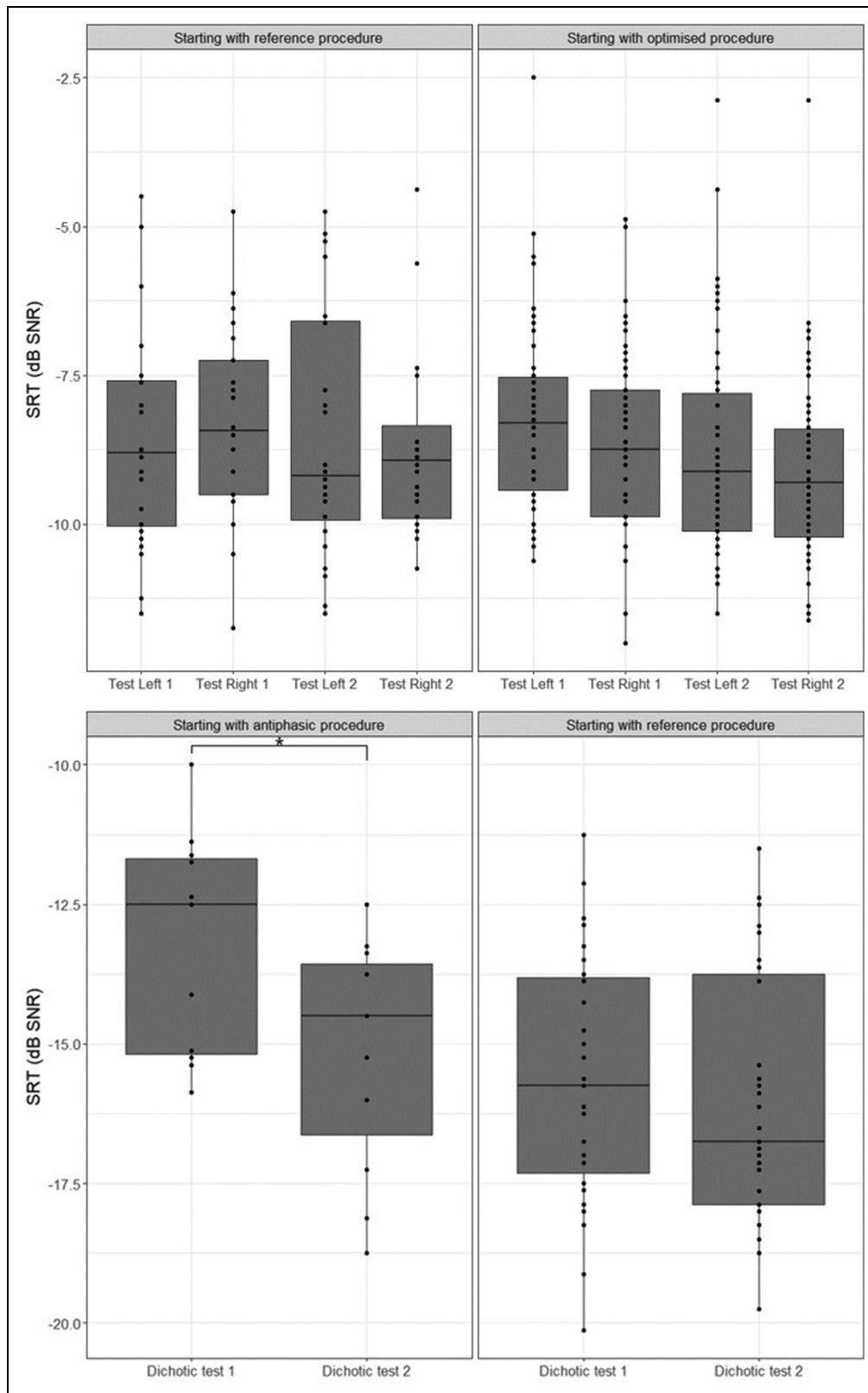
Moreover, Figure 8 shows the SRT-values obtained with the  $SEC_{REF}$  and the  $SEC_{APH}$  for the test and the retest. A decreasing trend in SRT-values is observed in this Figure for both the  $SEC_{REF}$  and  $SEC_{APH}$  in both sequences. However, for the  $SEC_{REF}$ , this decrease was not significant, and SRT-values in both ears were stable after the training phase. No significant differences were determined between the first and the second test when starting with the  $SEC_{REF}$  (left ear:  $Z = 1148$ ,  $p = 0.11$ ; right ear:  $Z = 1112$ ,  $p = 0.24$ ) and when doing the  $SEC_{REF}$  second (left ear:  $Z = 137$ ,  $p = 1$ ; right ear:  $Z = 187$ ,  $p = 1$ ). For the  $SEC_{APH}$ , when starting with the  $SEC_{APH}$ , the SRT-values of the second test were significantly better for the first test than the second one ( $Z = 54$ ,  $p = 0.02$ ). When doing the  $SEC_{APH}$  second, the SRT-values were stable after the training, and no significant differences were determined between the first and the second test ( $Z = 286$ ,  $p = 0.42$ ).

Additionally, the measurement error of  $SEC_{REF}$  was equal to that of  $SEC_{APH}$  and  $1.3 \pm 0.1$  dB SNR.

Lastly, for the  $SEC_{REF}$ , 4% scored an unreliable result in the first test (two left ear tests, one right ear test). For the  $SEC_{APH}$ , 5% of the children scored unreliably. Altogether, the total percentage of unreliable results was 5%. All unreliable tests showed an increasing trend in the adaptive staircase. Of the five children with unreliable results in their first test, only one had an unreliable result in the retest of that test. As the new procedures proceeded automatically, children who did not want to guess could proceed with the test without help from a test leader. No children reported hearing sounds inside the noise. No new problems that were not yet experienced during Experiment 1 were documented.

## General Discussion

The SEC had very good sensitivity and specificity to detect SNHL  $> 20$  dB (0.80–0.90) in young children at the age of school entry and is in line with values reported for the DTT, which is currently considered to be a reference screening test (Van den Borre et al., 2021). For CHL, very good sensitivity to pick up CHL  $> 40$  dB (0.80–0.90) and reasonable sensitivity to identify CHL  $> 20$  dB (0.60–0.70) were determined. Normally, suprathreshold tests are not sensitive



**Figure 8.** : SRTs for the SEC<sub>REF</sub> (top) and the SEC<sub>APH</sub> (bottom).

to CHL as CHL contains an attenuation component and no distortion component. However, as Denys et al. (2019) described, identifying the SEC's sounds is likely to rely on the temporal aspects of the stimuli used. The temporal processing is distorted when CHL is present (Balen et al., 2009; Bayat et al., 2017; Hartley & Moore, 2003; Moore et al., 2003). Therefore, this could be why CHL affected the identification of the sounds used in the SEC. Moreover, with a noise level fixed at 65 dB and an average SRT-value of approximately  $-9$  dB SNR, an audibility factor could have played a role in identifying the sounds for some children. The slightly lower sensitivity and specificity to mild CHL are less problematic as this type of hearing loss is mostly due to acute otitis media. Middle ear ventilation problems are common among preschool children and are usually self-limiting (Venekamp et al., 2014). Referring all children with acute CHL for extra examination would result in too many unnecessary diagnostics. On the other hand, identifying mild SNHL, which can be done with the SEC, is important as minimal SNHL can already affect hearing, speech understanding, and academic performance in daily life (Bess, 1985; Moore et al., 2019).

The duration of the original SEC was  $5 \text{ min } 50 \text{ s} \pm 1 \text{ min } 14 \text{ s}$ . This duration is comparable with the DTT done by children in the 5<sup>th</sup> grade of elementary school (Denys et al., 2018). The DTT has been successfully implemented for SHS in children in the 5<sup>th</sup> grade of elementary and 3<sup>rd</sup> grade of secondary school and has shown to be highly effective (Denys et al., 2018). However, the DTT was shown to be too difficult for younger children. Around 20% of the children at school entry could not perform a reliable DTT as a self-test (Denys et al., 2021). Before optimization, we found a similar number (17%) of unreliable test results for the SEC. Analyses of these unreliable test results showed that most unreliable results are due to attention and motivation loss instead of an overly high difficulty level.

With adapted test procedures, the number of unreliable results was reduced. With the SEC<sub>REF</sub>, only 4% of the young children scored an unreliable result in the first test. 5% of the children scored unreliable for the SEC<sub>APH</sub>. On average, only around 5% scored an unreliable result on a new, adaptive SEC procedure. This is a large reduction in unreliable tests compared to the SEC used in Experiment 1. However, it is impossible to tell whether this reduction is present purely due to the adaptations of the test. For example, the location of the test possibly could have affected the number of unreliable test results. In the first experiment, tests were done at home, in schools, and in the hospital, and in the second experiment, all tests were done at home. It is possible that children were less distracted and more focused at home. It is also possible that the slightly older population in Experiment 2 reduced the number of unreliable test results. Follow-up research is needed to clarify whether the number of unreliable tests stays low when the test is used in more variable circumstances.

With a test format that uses the reference procedure with 21 trials in the test phases, a duration of  $4 \text{ min } 35 \text{ s} \pm 52 \text{ s}$  was determined. With a procedure with a fixed SNR and a stopping rule, as described by Smits (2017), the duration was reduced to  $3 \text{ min } 51 \text{ s} \pm 1 \text{ min } 5$ , and with a procedure with dichotic stimuli, where both ears are tested together, the average duration was  $3 \text{ min} \pm 43 \text{ s}$ . The differences in duration were significant but were limited due to the acclimatization and training phase preceding the test phases. Moreover, the SRT-values of the antiphase procedure did not seem to be stable after a training phase with 21 trials, and a training effect seemed to be present between the first and the second antiphase test. When looking into other research that uses SPIN tests with dichotic stimuli, no reports could be found about the training time needed for a dichotic test (de Sousa et al., 2020; Smits et al., 2016; Wolmarans et al., 2021). However, a complete second training phase would increase the test by around 1 min, leaving the test duration at around 5 min, similar to the duration of the reference procedure with monaurally tested ears. Therefore, it would be interesting to research how much the training phase has to be increased to get stable SRTs, i.e., SRTs that do not change significantly with extra training afterwards.

The sensitivity and specificity were not explored in this research for the new, adapted procedures. Moreover, the cultural independence of the SEC was investigated in this research, but only with data from three different countries. Therefore, data on eventual cultural differences is limited. These unanswered questions are currently being studied in a large, international validation study. This follow-up research aims to research cultural independence to a larger extent. Moreover, it will help to gain knowledge on the sensitivity and specificity of the new procedures to detect childhood hearing loss.

## General Conclusion

High sensitivity and specificity of the SEC for CHL and SNHL were determined. After optimization of the sound- and noise material, the procedures, and the test interface, the number of unreliable tests could be reduced from 17% to only 5%. Therefore, the SEC has great potential to be used as an international hearing screening test for young children at school entry, giving countries and regions with little resources to develop and implement hearing screening programs. Consequently, the SEC can assist in the early detection of late-onset or acquired hearing loss in children internationally.



## Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the TBM-FWO Grant - Research Foundation-Flanders grant number (T002216N), Legaat Ghislaine Heylen (Tyberghein).

## ORCID iDs

Elie Van den Borre  <https://orcid.org/0000-0002-6352-0500>  
Sam Denys  <https://orcid.org/0000-0002-8692-4452>

## References

- Balen, S. A., Bretzke, L., Mottecy, C. M., Liebel, G., Boeno, M. R. M., & Gondim, L. M. A. (2009). Temporal resolution in children: Comparing normal hearing, conductive hearing loss and auditory processing disorder. *Brazilian Journal of Otorhinolaryngology*, *75*(1), 123–129. [https://doi.org/10.1016/S1808-8694\(15\)30843-0](https://doi.org/10.1016/S1808-8694(15)30843-0)
- Bayat, A., Farhadi, M., Emamdjomeh, H., Saki, N., Mirmomeni, G., & Rahim, F. (2017). Effect of conductive hearing loss on central auditory function. *Brazilian Journal of Otorhinolaryngology*, *83*(2), 137–141. <https://doi.org/10.1016/j.bjorl.2016.02.010>
- Bess, F. H. (1985). The minimally hearing-impaired child. *Ear and Hearing*, *6*(1), 43–47. doi:10.1097/00003446-198501000-00012
- Bess, F. H., Dodd-Murphy, J., & Parker, R. A. (1998). Children with minimal sensorineural hearing loss; prevalence, educational performance, and functional Status. *Ear & Hearing*, *19*(5), 339–354. doi:10.1097/00003446-199810000-00001
- Butcher, E., Dezateux, C., Cortina-Borja, M., & Knowles, R. L. (2019). Prevalence of permanent childhood hearing loss detected at the universal newborn hearing screen: Systematic review and metaanalysis. *PLoS One*, *14*(7), 1–21. <https://doi.org/10.1371/journal.pone.0219600>
- Danhauer, J. L., Chari, N. C. A., & Herman, G. (1977). Normal-hearing subjects' perception of speech through one-third octave bands. *The Journal of Auditory Research*, *17*(4), 275–281.
- Denys, S., de Laat, J., Dreschler, W., Hofmann, M., van Wieringen, A., & Wouters, J. (2019). Language-Independent hearing screening based on masked recognition of ecological sounds. *Trends in Hearing*, *23*, 1–16. <https://doi.org/10.1177/2331216519866566>
- Denys, S., Hofmann, M., Luts, H., Guérin, C., Keymeulen, A., van Hoeck, K., van Wieringen, A., Hoppenbrouwers, K., & Wouters, J. (2018). School-Age hearing screening based on speech-in-noise perception using the digit triplet test. *Ear and Hearing*, *39*(6), 1104–1115. doi:10.1097/AUD.0000000000000563
- Denys, S., Wouters, J., & Van Wieringen, A. (2021). The digit triplet test as a self-test for hearing screening at the age of school-entry the digit triplet test as a self-test for hearing screening at the age of school-entry. *International Journal of Audiology*, *0*(0), 1–8. <https://doi.org/10.1080/14992027.2021.1930204>
- de Sousa, K. C., Swanepoel, D. W., Moore, D. R., & Myburgh, H. C. (2020). Improving sensitivity of the digits-in-noise test using antiphase stimuli. *Ear & Hearing*, *41*(2), 442–450. doi:10.1097/AUD.0000000000000775
- Dettman, S. J., Pinder, D., Briggs, R. J. S., Dowell, R. C., & Leigh, J. R. (2007). Communication development in children who receive the cochlear implant younger than 12 months : Risks versus benefits. *Ear & Hearing*, *28*(2), 11–18. doi:10.1097/AUD.0b013e31803153f8
- Fletcher, H., & Galt, R. H. (1950). The perception of speech and its relation to telephony. *Journal of the Acoustical Society of America*, *22*(2), 89–151. <https://doi.org/10.1121/1.1906605>
- Fortnum, H. M., Summerfield, A. Q., Marshall, D. H., Davis, A. C., & Bamford, J. M. (2001). Prevalence of permanent childhood hearing impairment in the United Kingdom and implications for universal neonatal hearing screening: Questionnaire based ascertainment study. *BMJ*, *323*, 1–6. <https://doi.org/10.1136/bmj.323.7312.536>
- French, N. R., & Steinberg, J. C. (1947). Factors governing the intelligibility of speech sounds. *Journal of the Acoustical Society of America*, *19*(1), 90–119. <https://doi.org/10.1121/1.1916407>
- George, E. L. J., Festen, J. M., & Goverts, T. S. (2012). Effects of reverberation and masker fluctuations on binaural unmasking of speech. *The Journal of the Acoustical Society of America*, *132*(3), 1581–1591. <https://doi.org/10.1121/1.4740500>
- Greenwood, D. D. (1961). Auditory masking and the critical band. *The Journal of the Acoustical Society of America*, *33*(4), 484–502. <https://doi.org/10.1121/1.1908699>
- Hartley, D. E. H., & Moore, D. R. (2003). Effects of conductive hearing loss on temporal aspects of sound transmission through the ear. *Hearing Research*, *177*(1–2), 53–60. [https://doi.org/10.1016/S0378-5955\(02\)00797-9](https://doi.org/10.1016/S0378-5955(02)00797-9)
- Henry, B. A., Turner, C. W., & Behrens, A. (2005). Spectral peak resolution and speech recognition in quiet: Normal hearing, hearing impaired, and cochlear implant listeners. *The Journal of the Acoustical Society of America*, *118*(2), 1111–1121. <https://doi.org/10.1121/1.1944567>
- Holzinger, D., Hofer, J., Dall, M., & Fellingner, J. (2022). Multidimensional family-centred early intervention in children with hearing loss: A conceptual model. *Journal of Clinical Medicine*, *11*(6): 1548. <https://doi.org/10.3390/jcm11061548>
- Jansen, S., Luts, H., Wagener, K. C., Frachet, B., & Wouters, J. (2010). The French digit triplet test: A hearing screening tool for speech intelligibility in noise. *International Journal of Audiology*, *49*(5), 378–387. <https://doi.org/10.3109/14992020903431272>
- Koch, K. R. (2007). *Introduction to Bayesian statistics* (2nd ed.). Springer Science & Business Media.
- Kramer, S. E., Kapteyn, T. S., Festen, J. M., & Kramer, S. E. (2008). The self-reported handicapping effect of hearing disabilities. *International Journal of Audiology*, *37*(5), 302–312. <https://doi.org/10.3109/00206099809072984>
- Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *The Journal of the Acoustical Society of America*, *49*(2B), 467–477. <https://doi.org/10.1121/1.1912375>
- Mehra, S., & Eavey, R. D. (2009). The epidemiology of hearing impairment in the United States: Newborns, children, and adolescents. *Otolaryngology-Head and Neck Surgery*, *140*(4), 461–472. <https://doi.org/10.1016/j.otohns.2008.12.022>
- Moeller, M. P. (2000). Early intervention and language development in children who are deaf and hard of hearing. *Pediatrics*, *106*(3): e43–e43. <https://doi.org/10.1542/peds.106.3.e43>
- Moore, D. R., Hartley, D. E. H., & Hogan, S. C. M. (2003). Effects of otitis media with effusion (OME) on central auditory function. *International Journal of Pediatric Otorhinolaryngology*, *67*(SUPPL. 1), 63–67. <https://doi.org/10.1016/j.ijporl.2003.08.015>
- Moore, D. R., Zobay, O., & Ferguson, M. A. (2019). Minimal and mild hearing loss in children. *Ear and Hearing*, 1–732. <https://doi.org/10.1097/aud.0000000000000802>

- Myers, L. L., Letowski, T., & Abouchakra, K. (1996). Detection and recognition of octave-band sound effects. *Journal of American Academy of Audiology*, 7(5), 346–357.
- Neumann, K., Böttcher, P., & Euler, A. (2006). Effectiveness and efficiency of a universal newborn hearing screening in Germany. *Folia Phoniatrica et Logopaedica*, 58(6), 440–455. <https://doi.org/10.1159/000095004>
- Neumann, K., Mathmann, P., Chadha, S., Euler, H. A., & White, K. R. (2022). Newborn hearing screening benefits children, but global disparities persist. *Journal of Clinical Medicine*, 11(1), 271. doi:10.3390/jcm11010271
- Nolte, A., Coninx, F., Hess, M., Wiesner, T., Mu, F., Dudek, N., & Rohlf, A. (2016). Frequency-specific animal sound test (FAST) 4 : A valid method for hearing screening. *International Journal of Pediatric Otorhinolaryngology*, 81, 68–79. <https://doi.org/10.1016/j.ijporl.2015.12.004>
- Pimperton, H., Blythe, H., Kreppner, J., Mahon, M., Peacock, J. L., Stevenson, J., Terleksi, E., Worsfold, S., Yuen, H. M., & Kennedy, C. R. (2016). The impact of universal newborn hearing screening on long-term literacy outcomes: A prospective cohort study. *Archives of Disease in Childhood*, 101(1), 9–15. <https://doi.org/10.1136/archdischild-2014-307516>
- Plomp, R. (1978). Auditory handicap of hearing impairment and the limited benefit of hearing aids. *The Journal of the Acoustical Society of America*, 63(2), 533–549. <https://doi.org/10.1121/1.381753>
- Plomp, R. (1985). Relations between psychophysical data and speech perception for hearing-impaired subjects. II. *Journal of the Acoustical Society of America*, 78(4), 1261–1270. <https://doi.org/10.1121/1.392895>
- Porter, H., Sladen, D. P., Ampah, S. B., Rothpletz, A., & Bess, F. H. (2013). Developmental outcomes in early school-age children with minimal hearing loss. *American Journal of Audiology*, 22(2), 263–271. [https://doi.org/10.1044/1059-0889\(2013\)13-0013](https://doi.org/10.1044/1059-0889(2013)13-0013)
- Prieve, B. A., Schooling, T., Venedictov, R., & Franceschini, N. (2015). An evidence-based systematic review on the diagnostic accuracy of hearing screening instruments for preschool- and school-age children. *American Journal of Audiology*, 24(2), 250–267. <https://doi.org/10.1044/2015>
- RStudio (2020). *RStudio: Integrated Development for R. R studio*. PBC. <http://www.rstudio.com/>
- Smith, R. J. H., Bale, J. F., & White, K. R. (2005). Sensorineural hearing loss in children. *Lancet (London, England)*, 365(9462), 879–890. [https://doi.org/10.1016/S0140-6736\(05\)71047-3](https://doi.org/10.1016/S0140-6736(05)71047-3)
- Smits, C. (2017). Improving the efficiency of speech-in-noise hearing screening tests. *Ear & Hearing*, 38(6), e386–e388. <https://doi.org/10.1097/AUD.0000000000000446>
- Smits, C., & Houtgast, T. (2005). Results from the Dutch speech-in-noise screening test by telephone. *Ear & Hearing*, 26(1), 89–95. <https://doi.org/10.1097/00003446-200502000-00008>
- Smits, C., Kapteyn, T. S., & Houtgast, T. (2004). Development and validation of an automatic speech-in-noise screening test by telephone. *International Journal of Audiology*, 43(1), 15–28. <https://doi.org/10.1080/14992020400050004>
- Smits, C., Watson, C. S., Kidd, G. R., Moore, D. R., & Goverts, S. T. (2016). A comparison between the Dutch and American-English digits-in-noise (DIN) tests in normal-hearing listeners. *International Journal of Audiology*, 55(6), 358–365. <https://doi.org/10.3109/14992027.2015.1137362>
- Van den Borre, E., Denys, S., Van Wieringen, A., & Wouters, J. (2021). The digit triplet test : A scoping review. *International Journal of Audiology*, 1–18. <https://doi.org/10.1080/14992027.2021.1902579>
- Venekamp, R. P., Damoiseaux, R. A. M. J., & Schilder, A. G. M. (2014). Acute otitis media in children. *BMJ Clinical Evidence*, 2014, 1–21. <https://doi.org/10.1136/bmj.m4238>
- Watkin, P., McCann, D., Law, C., Mullee, M., Petrou, S., Stevenson, J., Worsfold, S., Ho, M. Y., & Kennedy, C. (2007). Language ability in children with permanent hearing impairment: The influence of early management and family participation. *Pediatrics*, 120(3), e694–e701. <https://doi.org/10.1542/peds.2006-2116>
- Winiger, A., Alexander, J., & Diefendorf, A. (2016). Minimal hearing loss: From a failure-based approach to evidence-based practice. *American Journal of Audiology*, 25(3), 1–15. <https://doi.org/10.1044/2016>
- Wolmarans, J., De Sousa, K. C., Frisby, C., Mahomed-Asmail, F., Smits, C., Moore, D. R., & Swanepoel, D. W. (2021). Speech recognition in noise using binaural diotic and antiphase digits-in-noise in children: Maturation and self-test validity. *Journal of the American Academy of Audiology*, 32(5), 315–323. <https://doi.org/10.1055/s-0041-1727274>
- World Health Organization (2021). Hearing screening: Considerations for implementation. *Licence: CC BY-NC-SA 3.0 IGO*.
- Wouters, J., de Laat, J. A. P. M., Denys, S., Dreschler, W. A., Hoffman, M., & van Wieringen, A. (2017). EFAS working group on school age hearing screening. *EFAS Working Group on School Age Hearing Screening* [Unpublished manuscript].
- Yong, M., Panth, N., McMahon, C. M., Thorne, P. R., & Emmett, S. D. (2020). How the world's children hear: A narrative review of school hearing screening programs globally. *OTO Open*, 4(2): 2473974X2092358. <https://doi.org/10.1177/2473974X20923580>
- Yoshinaga-itano, C., Coulter, D., & Thomson, V. (2001). Developmental outcomes of children with hearing loss born in Colorado hospitals with and without universal newborn hearing screening programs. *Semin Neonatal*, 6(6), 521–529. <https://doi.org/10.1053/siny.2001.0075>
- Yoshinaga-Itano, C., Sedey, A. L., Coulter, D. K., & Mehl, A. L. (1998). Language of early- and later-identified children with hearing loss. *Pediatrics*, 102(5), 1161–1171. <https://doi.org/10.1542/peds.102.5.1161>