Speech Recognition Abilities in Normal-Hearing Children 4 to 12 Years of Age in Stationary and Interrupted Noise

Wiepke J. A. Koopmans,[1,2] S. Theo Goverts,[1] and Cas Smits[1]

Objectives: The main purpose of this study was to examine developmental effects for speech recognition in noise abilities for normal-hearing children in several listening conditions, relevant for daily life. Our aim was to study the auditory component in these listening abilities by using a test that was designed to minimize the dependency on nonauditory factors, the digits-in-noise (DIN) test. Secondary aims were to examine the feasibility of the DIN test for children, and to establish age-dependent normative data for diotic and dichotic listening conditions in both stationary and interrupted noise.

Design: In experiment 1, a newly designed pediatric DIN (pDIN) test was compared with the standard DIN test. Major differences with the DIN test are that the pDIN test uses 79% correct instead of 50% correct as a target point, single digits (except 0) instead of triplets, and animations in the test procedure. In this experiment, 43 normal-hearing subjects between 4 and 12 years of age and 10 adult subjects participated. The authors measured the monaural speech reception threshold for both DIN test and pDIN test using headphones. Experiment 2 used the standard DIN test to measure speech reception thresholds in noise in 112 normal-hearing children between 4 and 12 years of age and 33 adults. The DIN test was applied using headphones in stationary and interrupted noise, and in diotic and dichotic conditions, to study also binaural unmasking and the benefit of listening in the gaps.

Results: Most children could reliably do both pDIN test and DIN test, and measurement errors for the pDIN test were comparable between children and adults. There was no significant difference between the score for the pDIN test and that of the DIN test. Speech recognition scores increase with age for all conditions tested, and performance is adult-like by 10 to 12 years of age in stationary noise but not interrupted noise. The youngest, 4-year-old children have speech reception thresholds 3 to 7 dB less favorable than adults, depending on test conditions. The authors found significant age effects on binaural unmasking and fluctuating masker benefit, even after correction for the lower baseline speech reception threshold of adults in stationary noise.

Conclusions: Speech recognition in noise abilities develop well into adolescence, and young children need a more favorable signal-to-noise ratio than adults for all listening conditions. Speech recognition abilities in children in stationary and interrupted noise can accurately and reliably be tested using the DIN test. A pediatric version of the test was shown to be unnecessary. Normative data were established for the DIN test in stationary and fluctuating maskers, and in diotic and dichotic conditions. The DIN test can thus be used to test speech recognition abilities for normal-hearing children from the age of 4 years and older.

Key words: Age factors, Binaural unmasking, Child, Fluctuating masker benefit, Hearing tests, Speech intelligibility, Speech-in-noise recognition, Speech perception, Speech recognition abilities, Speech reception threshold test.

INTRODUCTION

Young children spend many hours a day in complex acoustic environments with noise and reverberation such as kindergarten and school. In these demanding listening situations, they have to communicate with their parents, teachers, and other children. Previous studies have shown that children have more difficulty than adults with recognizing speech in noisy situations (Crandell 1993; Hall et al. 2002), and that speech recognition abilities in noise develop at least to the age of 10 to 12 years (Buss et al. 2006; Hall et al. 2004; Holder et al. 2016; Vaillancourt et al. 2008). Children's reduced speech recognition abilities in noise may affect how well they learn in a noisy classroom, through both formal education and incidental learning. On top of this developmental effect, the ability to recognize speech in noise can be strongly reduced by hearing loss (Ching et al. 2017), which makes daily-life listening conditions often critical for children with hearing impairment. To quantify the consequences of hearing loss in children with hearing impairment, it is important to relate the outcomes of hearing assessment to those of their normal-hearing peers. Hence, it is important and clinically relevant to know how speech recognition abilities in noise of normal-hearing children develop with age.

Listening in an acoustically demanding situation involves combining the two different, but related noise-corrupted speech fragments from both ears. Children's speech recognition abilities therefore depend on their ability to separate speech from noise, to benefit from fluctuations in the background noise, and to benefit from binaural cues. Previous studies have shown that children's test performance on speech-in-noise tests improves with age (Corbin et al. 2016; Elliott 1979; Hall et al. 2002). In stationary speech-shaped noise, most children achieved adult-like performance by 10 years of age or later (Corbin et al. 2016; Elliott 1979; Hall et al. 2002; Holder et al. 2016; Neuman et al. 2010; Nishi et al. 2010; Wilson et al. 2010). Other masker types result in larger and more prolonged age effects. For example, Leibold and Buss (2013) found that the developmental trajectory depends on the masker type, and found a more prolonged developmental time course for consonant detection in two-talker babble than in speech-shaped noise masker. Corbin et al. (2016) found a similar prolonged developmental time course for word detection in two-talker babble compared with a speech-shaped masker. They hypothesized that masked speech recognition may rely on mature executive function to a greater extent in two-talker speech (informational masking) than in speech-shaped noise (energetic masking), and that this places a greater cognitive load on the child. This shows that the development of speech recognition in noise abilities has been

[1]Department of Otolaryngology-Head and Neck Surgery, Section Ear and Hearing, VU University Medical Center, Amsterdam, The Netherlands; and [2]NSDSK, Dutch Foundation for the Deaf and Hard of Hearing Child, Amsterdam, The Netherlands.

explained in terms of both auditory and nonauditory factors, and that the exact developmental time course likely depends on test procedures and masker types.

The effects of fluctuations in the noise on the speech recognition abilities of children have been studied by Stuart (2005, 2008), Hall et al. (2012), and Buss et al. (2016). They found that the test performance improves with age, and that children under the age of 11 to 14 years old need a more favorable signal to noise ratio (SNR) to perform as well as adults. Stuart (2008) measured performance in five groups of children (6–7, 8–9, 10–11, 12–13, and 14–15 years) and in adults, and found that the fluctuating masker benefit (FMB; i.e., the release of masking because of the interruptions in the noise) for the children was not significantly different from that of adults. He suggested that school-age children have an inherent poorer central processing efficiency, rather than poorer temporal resolution. Thus, children benefit from listening in the "gaps'" but their ability to recognize speech in noise is limited by ongoing maturation of the auditory system and their developing language and attention skills. Hall et al. (2012) tested the effect of temporally modulated maskers (100% sinusoidal modulation at a rate of 10 Hz) on speech recognition scores of children 4.6 to 11.1 years of age. They found a significant developmental effect on masking release related to the temporal modulations in the noise. In a later study, Buss et al. (2016) found similar age effects on masking release by a modulated masker in a four-alternative forced-choice response context. In both studies, the authors speculated that young children are relatively poor in the ability to piece together sparse "glimpses" of speech. They also noted that the observed developmental effect might be related to the relatively high SNR associated with baseline-masked thresholds in younger children (Bernstein & Grant 2009; Smits & Festen 2013). The origin of the observed child–adult differences in FMB is therefore not entirely clear.

The development of binaural hearing abilities has been studied in various ways. Binaural hearing abilities of adults are often assessed using headphones with binaural intelligibility level difference tests (Johansson & Arlinger 2002; Licklider 1948). These tests use the masking level difference when the speech is phase shifted between the right and left ear compared with the homophasic condition. The binaural unmasking [BU; i.e., the difference in speech reception threshold (SRT) between diotic (N0S0) and dichotic (N0Sπ) presentation] can amount to 7 dB SNR for adults (Johansson & Arlinger 2002). BU has been explored in tests with children, but typically only in tone discrimination tasks. Moore et al. (2011) did not find a significant change in masking level difference with age, but they also state that the small number of children examined in their study may have contributed to the nonsignificant result. Several other groups have studied children's ability to benefit from spatial and binaural hearing when target speech and competing noise are spatially separated in a sound field. There is no consensus on how this spatial release of masking (SRM) develops with age (as reviewed by Yuen & Yuan 2014). Some studies report that SRM does not improve with age and becomes adult-like at a young age (Ching et al. 2011; Garadat & Litovsky 2007; Litovsky 2005; Murphy et al. 2011). For example, Litovsky (2005) used target speech that was presented from the front, while speech or modulated speech-shaped noise competitors were either in front or on the right at 90°. She found that SRM was similar in the two age groups (children 4 to 7 years of age and adults), and even greater in children in one condition tested. Her findings suggested that young children are already able to utilize spatial and/or head shadow cues to segregate sounds in noisy environments. By contrast, other studies report that SRM improves with age and that it takes much longer to reach adult-like performance (Cameron et al. 2009; Cameron & Dillon 2007; Vaillancourt et al. 2008; Van Deun et al. 2010; Yuen & Yuan 2014). For example, Van Deun et al. (2010) used a speech test with digits in noise to measure speech perception benefits in normal-hearing children between 4 and 8 years of age and normal-hearing adults. They measured SRM, head shadow effects, summation effects, and squelch and found that only SRM was influenced by age. Yuen and Yuan (2014) revisited the research question on whether the development of SRM is completed early or late in children. They hypothesized that there is a much longer maturational time for SRM than suggested by other studies (Ching et al. 2011; Garadat & Litovsky 2007; Litovsky 2005; Murphy et al. 2011) because of the ongoing maturation of the auditory system. They performed SRM testing with children (4 to 9 years of age) and adults and found that SRM improves significantly with age. A robust regression of 0.1 to 0.15 dB SRM improvement per month was observed for two different test materials.

To summarize, a substantial body of literature demonstrates lower speech recognition scores in noise for children than for adults in various conditions that are relevant for everyday listening. The developmental time course depends on test materials and masker types used. The origin of the observed age dependencies for different speech recognition in noise abilities remains to be explained, but the literature suggests that auditory and nonauditory factors play a role.

Assessing and Interpreting Speech Recognition Abilities in Children

When assessing and interpreting speech recognition test scores in children, one has to consider various factors that can influence the conclusions based on the test: the test should measure the same speech recognition ability for children and adults; age-dependent normative data should be available; and the effect of the baseline SNR at which the masking release is estimated (Smits & Festen 2011) has to be taken into account.

First, standard speech-in-noise tests designed for adults may not be suitable for children because of the procedures and materials that are used. For example, the standard Dutch speech-in-noise tests [sentence speech-in-noise tests from Plomp & Mimpen (1979) and Versfeld et al. (2000)] are not suitable for children under 12 years of age because of the language competency required to complete the test. Mendel (2008) points out that test performance can be influenced by the child's vocabulary, language competency, and cognitive abilities. These nonauditory (or top–down) processes are developing during childhood. It is therefore difficult to discriminate between the developmental aspects of purely auditory speech recognition abilities (or bottom–up) and top–down processes. Moore et al. (2011) point out that children's reduced performance on auditory tasks may primarily be due to nonsensory factors. The poorer test performance of children is often explained in terms of "elevated internal noise" or "poor processing efficiency," although these concepts are ill defined. Because both auditory (bottom–up) and nonauditory (top–down) factors are developing in children,

<zdoi; 10.1097/AUD.0000000000000569>

age-appropriate tests are needed. Either target speech material, competitor noise, or the adaptive test procedure must be modified to meet the needs of children. Although Mendel (2008) gives a guide how to design test materials to be age appropriate, it is unclear to what extent the test scores and findings are still impacted by the child's cognitive abilities, attention skills, and linguistic proficiency. It is important to minimize the effect of these nonsensory factors on the test from purely auditory factors to primarily test the auditory bottom–up component of speech recognition in noise (Smits et al. 2013).

Second, to relate the test results of a child to their peers, age-dependent normative data should be available for the speech-in-noise test. Establishing normative data is quite an effort because many children from different age groups have to be tested for a clear understanding of the age-dependent mean test score and confidence interval. For practical reasons, there are some advantages to acquiring these normative data by using headphones, rather than in a free field. Free-field tests have to be conducted in a sound booth at a clinic and are susceptible to the child's head movements, and variation in position relative to the loudspeaker. When using acoustically isolated headphones, children's head movements have little effect on the test result and the tests may be performed outside the test booth, such as at school. This greatly facilitates the recruitment of an adequate sample size for each age group involved.

Finally, when considering the effect of age on masking release (FMB, SRM, or BU), one must realize that the amount of unmasking may depend on the baseline SNR at which unmasking is estimated (Bernstein & Grant 2009; Oxenham & Simonson 2009; Smits & Festen 2011). The slope of the speech recognition function is, in general, shallower in fluctuating noise than in steady-state noise (Smits & Festen 2013). Because of slope differences between the speech recognition function for the baseline stationary noise condition and speech recognition functions for other listening conditions (e.g., fluctuating maskers or spatially separated maskers), the FMB, SRM, or BU (i.e., the difference between these functions expressed in dB SNR) may depend on the baseline SNR. Thus, the FMB, SRM, or BU may be lower for children than for adults because children need a higher SNR in the baseline condition than adults. This effect is often overlooked in studies reported in the literature, and could potentially explain part of the reported age dependence of FMB and SRM. Therefore, when studying the effect of age on FMB and SRM, it is important to take the baseline SNR into account, and ideally, performance should be measured across the psychometric function. In summary, the assessment of speech recognition abilities in children could be facilitated by a speech recognition test that is applicable to children and adults; for which age-dependent normative data are available; and for which the effect of baseline SNR on masking release can be taken into account.

Digits-in-Noise Test

Smits et al. (2013) developed a digits-in-noise (DIN) test that was designed to measure primarily the auditory, bottom–up, speech recognition abilities in noise. The DIN test measures the SRT (i.e., the SNR corresponding to 50% correct recognition) for digit–triplets in long-term average speech spectrum (LTASS) noise. Smits et al. (2013) validated the test for adults and found that after a practice run, there is no residual

learning effect. There is a high correlation ($r = 0.96$) with SRT scores obtained with the standard sentences speech-in-noise test (Plomp & Mimpen 1979). Because of the steep speech recognition function, the DIN test has a small measurement error of only 0.7 dB (Smits et al. 2013). Because test scores on the DIN test hardly (<1 dB) depend on linguistic abilities (Kaandorp et al. 2016), the DIN test can be used in virtually the entire population of adults with hearing loss, from normal-hearing listeners to listeners with severe to profound hearing losses and cochlear implant recipients (Kaandorp et al. 2015). The DIN test has not been used in children before.

Aims of the Study

The primary purpose of this study was to examine developmental effects on speech recognition in noise abilities for normal-hearing children 4 to 12 years of age in stationary and interrupted noise. We used the DIN test to minimize the dependency on nonauditory factors. In Experiment 1, results on the DIN test are compared with test results on an adapted version of the DIN test, the pediatric DIN (pDIN) test that was designed to rule out contribution of specific nonauditory factors that might influence test performance for the youngest children. The factors are related to the test procedures and test materials used in the DIN test. For example, the DIN test uses a task that requires the reproduction of three digits. Normative data from the digit span tests in the Wechsler Intelligence Scale for Children show that the digits span increases with age and that 1.5% of the 6-year-old children cannot recall three digits in the forward direction (Wechsler 2004). This suggests that a small fraction of the 6-year-old and probably even larger fractions of 4- and 5-year-old children do not have the auditory memory to do the DIN test reliably. The pDIN test uses the same speech tokens as the DIN test, with the digit "0" omitted and in a single-digit paradigm, to circumvent this issue. Other modifications were made to simplify the test and make it appealing even to the youngest children (see Experiment 1 for details).

In Experiment 2, age-dependent normative data in normal-hearing children 4 to 12 years of age were established for the DIN test under headphones for N0S0 and N0Sπ listening conditions in both stationary and interrupted noise. We analyzed FMB and BU with respect to the baseline SNRs to determine "true" developmental effects of speech recognition in noise abilities, and describe the developmental time course for these effects.

EXPERIMENT I: A COMPARISON BETWEEN THE PDIN TEST AND THE STANDARD DIN TEST

A pDIN test was developed to simplify the test and make it appealing even to the youngest children. Results of the pDIN test were compared with those of the DIN test to find out which factors influence test performance.

Materials and Methods

Subjects • In this study, 43 native Dutch-speaking, normal-hearing children (22 male and 21 female) between 4 and 13 years of age participated. They were recruited from a local primary school. Their parents or caregivers, and children of 12 years of age and older gave their written informed consent. To determine adult reference data, 10 native Dutch-speaking

(1 male, 9 female), normal-hearing adults between 18 and 33 years of age participated in the study. Normal hearing was defined as air conduction thresholds equal to or better than 20 dB HL for all octave frequencies from 0.25 to 8 kHz in the test ear. All subjects had normal (type A) tympanograms.

DIN Test Stimuli and Test Procedure • The speech material and masking noise used in the DIN test are described elsewhere in detail (Smits et al. 2013). Briefly, the DIN test uses a set of 120 unique digit–triplet combinations constructed from the digits 0 to 9 uttered by a male speaker, separated by short (150 msec), silent intervals. Each triplet stimulus started and ended with 500 msec of silence. All the silent intervals were enlarged or reduced with an interval chosen at random between +50 and −50 msec to add uncertainty to the listening task. The stimulus was mixed with LTASS masking noise to achieve the desired SNR. The noise started and ended with a 100 msec raised cosine ramp. The duration of the triplet-in-noise files ranged from 2.8 to 3.1 sec. Signal and noise were presented at a fixed overall level of 65 dBA. The SNR was varied adaptively following the standard one-up one-down procedure with a step size of 2 dB SNR. The first stimulus was presented at a favorable SNR of 6 to 8 dB above the expected SRT, and 24 triplets were presented. The SNR for triplet 25 was calculated but not presented. The DIN SRT was calculated by taking the average SNR of trial 5 to 25.

pDIN Test Stimuli and Test Procedure • The pDIN test uses the single digits 1 to 9 from the same speech material as the DIN test, but in a single-digit format. The digit 0 was omitted from the test because the concept of zero needs longer to develop in children (Wellman & Miller 1986). Each digit stimulus started and ended with 500 msec of silence, enlarged or reduced with an interval chosen at random between +50 and −50 msec to add uncertainty to the listening task. The stimulus was mixed with LTASS masking noise to achieve the desired SNR. Each presentation started and ended with a 100 msec raised cosine ramp. The test first presented the digits 1 to 9 in random order in quiet to find out if the child could reproduce each digit correctly. If the child did not repeat a particular digit correctly, it was presented a second time. The digit was automatically omitted from the test if the child responded incorrectly a second time. Then the noise was introduced with an animation, depicting a scientist that builds a "noise-machine," Next, the adaptive test procedure started. Signal and noise were presented at a fixed overall level of 65 dBA. The SNR was varied adaptively, with a weighted up–down procedure (Kaernbach 1991). The step size for trial 1 to 4 was 3 dB, to approach the SRT quickly. Step sizes for trial 5 to 24 were 0.67 dB down and 2.57 dB up, such that the 79.4% point of the psychometric function was targeted. By choosing this target point, the pDIN SRT for the pDIN test and the DIN SRT for the DIN test correspond to the same SNR, and can theoretically be compared. The SRT for the DIN test is defined as the SNR where 50% of the triplets are reproduced correctly. The triplet consists of concatenated digits without prosody or coarticulation. This means that the probability of reproducing an individual digit (P_{digit}) correctly at this SNR is statistically independent from the other digits in the triplet (Smits & Houtgast 2006). The probability of reproducing a triplet correct ($P_{triplet}$) is then given by a simple product of the probabilities of the individual digits in this triplet: $P_{triplet} = P_{digit}^3 = 0.5$. Hence, $P_{digit} = \sqrt[3]{0.5} = 0.794$, which means 79.4% of the single digits are reproduced correctly at this target point.

Lists of 24 digits were presented, such that each digit 1 to 9 was presented at least twice and at most three times, and that consecutive digits were never the same. The SNR for triplet 25 was calculated but not presented. The child repeated the digits, and the experimenter recorded the response in the computer program. The pDIN SRT was calculated by taking the average SNR of presentations 5 to 25. Dummy presentations of digits at a favorable SNR (+5 dB from the current estimated SRT) were presented after animations every six trials to keep the child motivated and alert. The response to the dummy presentation was not used for calculating the SRT.

Setup • All tests with children were carried out in a quiet office room at the local primary school. Air conduction pure-tone audiograms were measured with a portable clinical audiometer (Noordwijk, The Netherlands: Decos Technology) and Sennheiser HDA 200 headphones (Wedemark, Germany: Sennheiser electronic GmbH & Co. KG). Custom software (Austin, Texas: Delphi Embarcadero Technologies) was developed for the pDIN test and DIN test. It presents speech and noise stimuli at a defined SNR, records and judges the response, adjusts the SNR, and stores the results in a database. All stimuli were presented monaurally through Sennheiser HDA 200 headphones, connected to a digital sound card (Soundblaster Audigy; Dublin, Ireland: Creative Technology Ltd) and a laptop. All tests with adults were carried out with the same equipment, in a standard, quiet office room at VU University Medical Center, Amsterdam.

Overall Test Procedure • Each session started with a pDIN test practice run to familiarize the child with the task and eliminate procedural learning effects. Next, the child performed a pDIN test and retest, and finally, a DIN practice run followed by a single DIN test. In the initial test phase, only pDIN test measurements were performed because we expected that the DIN test would be too difficult for the younger children. However, during the experiment, it became apparent that testing with the DIN test was feasible for almost all of the children and the DIN test was administered in all children from then on. Therefore, the DIN test was not performed by all subjects. Finally, the experimenter determined the child's pure-tone audiogram. Testing sessions took 20 to 30 min per subject. The procedure for testing adults was similar, except that animations and dummy trials were not presented in the pDIN test. The study was approved by the VU University Medical Centre Medical Ethical Committee.

Results

pDIN and DIN Test Feasibility • All children of 4 years of age and older could repeat the numbers 1 to 9 in quiet, thus no digits were omitted from the test. The administration of multiple tests in a single test session was possible for all children. All 43 children did a pDIN test practice run and a test, and 41 did a retest. For the children who were tested with the DIN test (N = 35), only 1 did not complete the test after the practice run. Thus, 34 from 35 children performed five SRT tests in one session and only 1 completed four SRT tests. A single test took approximately 3 min. The task of reproducing either single digits or triplets in noise could be performed even by the youngest children.

The mean SRT scores and the standard error of the mean (SEM; derived from the test–retest differences) for the pDIN test for different age groups are summarized in Table 1. The measurement error for the pDIN test, represented by the SEM, was approximately 1 dB for all age groups (range, 0.9 to 1.1 dB).

<zdoi; 10.1097/AUD.0000000000000569>

TABLE 1. Group mean monaural SRT scores for the DIN and pDIN tests for different age groups, measurement error for the pDIN test, and difference between pDIN and DIN SRT

Age Group	N (pDIN; DIN)	pDIN SRT* (dB SNR)	DIN SRT* (dB SNR)	pDIN SEM (dB)	pDIN–DIN (dB)
4–5 yrs	23; 17	−5.8 ± 1.2	−5.8 ± 1.0	1.1	−0.29 [t(16) = −1.1; p = 0.3]
6–8 yrs	14; 14	−7.0 ± 1.2	−7.4 ± 0.9	0.9	0.39 [t(13) = 1.2; p = 0.3]
9–12 yrs	6; 3	−8.6 ± 1.4	−8.4 ± 0.3	1.0	0.07 [t(2) = 0.06; p = 0.96]
Adult	10; 10	−8.9 ± 1.1	−9.5 ± 0.5	1.0	0.59 [t(9) = 2.0; p = 0.07]

*Mean ± SD.
DIN, digits-in-noise; pDIN, pediatric digits-in-noise; SEM, standard error of the mean; SNR, signal to noise ratio; SRT, speech reception threshold.

A repeated-measures analysis of variance (ANOVA) was conducted to compare the effects of age group and learning (test versus retest) on the pDIN SRT. There was a significant effect of age group [$F(3,47) = 22.2$; $p < 0.001$]. There was neither a significant effect of learning [$F(1,47) = 0.00$] nor a significant interaction between age group and learning [$F(3,47) = 0.55$; $p = 0.562$]. These results suggest that SRTs change with age and that there is no residual learning effect between the first and second test after the practice run.

When the digit 0 was presented as part of a triplet in the DIN test, children 4 to 5 years of age reproduced this digit correctly 82% of the trials across the different SNRs presented. Adults reproduced this digit correctly 74% of the trials across the different SNRs presented. This percentage was in the same range of percentages correct for the digits 1 to 9 (67 to 88% for the children 4 to 5 years of age and 64 to 91% for the adults). This finding shows that the digit 0 can be used in the DIN test for young children of 4 to 5 years of age and older.

Age Dependency of DIN SRT and pDIN SRT • Figure 1 shows the monaural SRT measured with the DIN test and pDIN test as a function of age. The individual test scores for each child are shown in a scatter plot (retest scores are not shown). The box plot represents the results for the adult group, with median group SRT (horizontal line), 25th and 75th percentile SRT (box ends), 10th and 90th percentile SRT (whiskers), and outliers (open circles). The thick line is an exponential fit to the data from the children. The regression equations are shown in Figure 1. The intersubject variance (spread in SRT values within the group) was different for both tests, for children and adults.

After correction for the age-dependent group mean SRT, the standard deviation for the DIN test and pDIN test was 0.8 and 1.2 dB for children, and 0.34 and 1.1 dB for adults, respectively. For children, this observed reduction in variance is consistent with the three times greater number of presentations in the DIN test relative to the pDIN test, derived from model calculations by Smits and Houtgast (2006). For adults, the reduction in variance is larger than predicted.

Equivalence of DIN SRT and pDIN SRT • Figure 2A shows the DIN SRT as a function of the pDIN SRT. Given the relatively small range in SRT values, there is still a reasonably strong, positive correlation between the two. Pearson correlation coefficient was 0.74 for a single pDIN test and DIN test, and 0.85 when the average SRT of test and retest for the pDIN test was used.

Figure 2B shows the difference between DIN SRT and pDIN SRT as a function of age. Linear regression gave nonsignificant values for both offset (−0.42 dB; $p = 0.22$) and slope (0.06 dB/yr; $p = 0.07$). A paired t test of DIN SRT and pDIN SRT showed that there is no significant difference between pDIN and DIN (pDIN SRT–DIN SRT = 0.15 dB; $p = 0.4$). These results indicate that the pDIN test and DIN test do not yield a significantly different SRT value for children and adults.

EXPERIMENT II: CHILDREN'S SPEECH RECOGNITION ABILITIES IN STATIONARY AND INTERRUPTED NOISE

Experiment I showed that both the pDIN test and DIN test can reliably be performed by normal-hearing children between
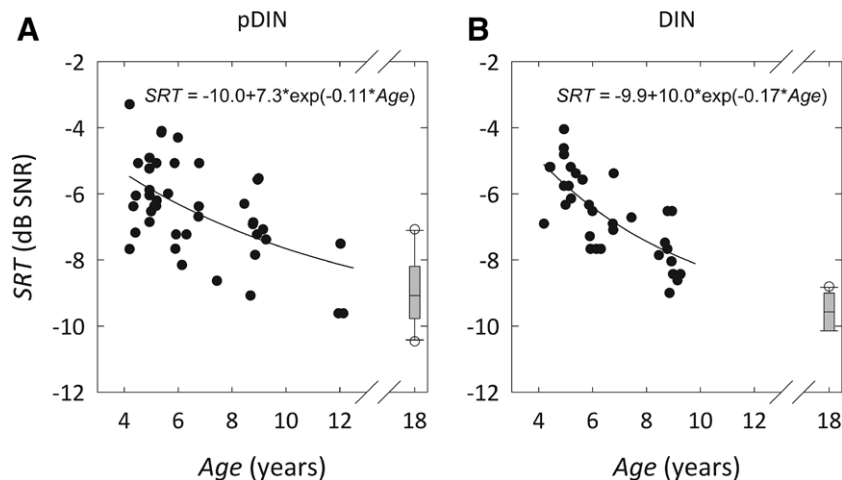

Fig. 1. Age Dependency of DIN SRT and pDIN SRT for monaural presentation. (A) pDIN SRT vs. age. (B) DIN SRT vs. age. The thick line represents an exponential fit to the data. DIN indicates digits-in-noise; pDIN, pediatric digits-in-noise; SNR, signal to noise ratio; SRT, speech reception threshold.
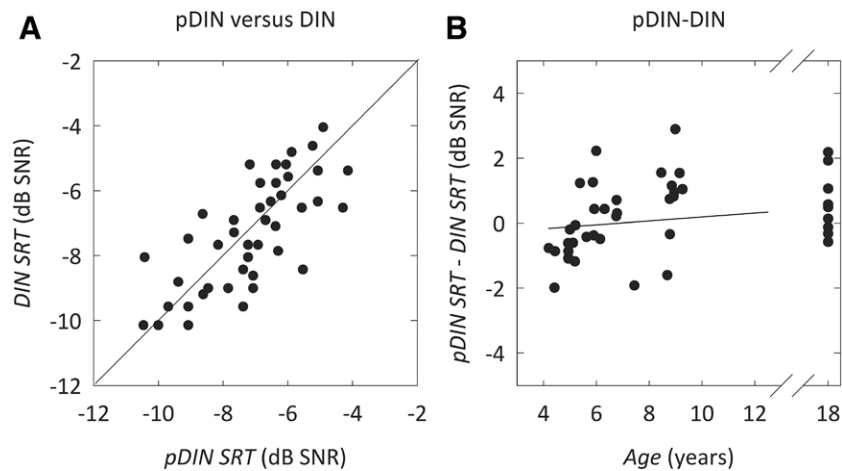
<zdoi; 10.1097/AUD.0000000000000569>

Fig. 2. Equivalence of DIN SRT and pDIN SRT. A, DIN SRT vs. pDIN SRT. Pearson $r = 0.74$ for a single test and Pearson $r = 0.80$ for the average of test and retest. The thick line represents the equal-SRT line. B, The difference in pDIN SRT–DIN SRT as a function of age. The slope and offset of the linear fit (represented by the thick line) are not significantly different from zero. DIN indicates digits-in-noise; pDIN, pediatric digits-in-noise; SNR, signal to noise ratio; SRT, speech reception threshold.

4 and 12 years of age and that they result in similar SRTs. Because the DIN SRT showed a smaller intersubject variance than the pDIN SRT, it was decided to further use the DIN test in Experiment II. The aim of Experiment II was to investigate speech recognition abilities in stationary and interrupted noise, and to determine the developmental time course of the benefit from BU and fluctuating maskers in children. BU was investigated by comparing N0S0 with N0Sπ presentation, while the FMB was measured by comparing SRTs in stationary noise with SRTs in interrupted noise. Finally, the combined effect of N0Sπ presentation and interrupted noise was studied.

Materials and Methods

Subjects • One hundred twelve native Dutch-speaking, normal-hearing children (57 male, 55 female) between 3 and 12 years of age recruited from a local primary school or through informal connections participated in this experiment. Their parents or caregivers gave their written informed consent. Thirty-three native Dutch-speaking (8 male, 25 female), normal-hearing adults between 18 and 30 years of age were recruited from the local university or informally to determine adult reference values for the test conditions. All children and adults had pure-tone air conduction threshold equal to or better than 20 dB HL at all octave frequencies from 250 to 8000 Hz, except 3 adult subjects with a higher threshold at one frequency in one ear.

Stimuli • The digit–triplets and stationary noise from the DIN test as in Experiment 1 were used. Interrupted noise was constructed by modulating the stationary noise with a 50% duty cycle, 16 Hz square wave. The modulation depth was 15 dB. Stimuli were presented either diotically or dichotically. In the N0S0 conditions, the noise and stimuli were identical in both left and right channels. In N0Sπ conditions, the noise was presented in phase in both channels and the speech out of phase by inverting one of the channels. With these stimuli, four different conditions could be tested: N0S0 presentation using stationary (N0S0$_{stat.}$) or fluctuating noise (N0S0$_{fluct.}$), and N0Sπ presentation using stationary (N0Sπ$_{stat.}$) or fluctuating noise (N0Sπ$_{fluct.}$).

Setup • The children were tested in a quiet office room at the local primary school. Tests with adults were carried out in a

quiet office room or in a double-walled sound proof booth at VU University Medical Center, or in a quiet room at their home. The equipment for the pure-tone audiometry and for the DIN test was the same as in Experiment I.

Procedure • The experimenter first checked if the child was familiar with all the digits 0 to 9 by asking the child to count from 0 to 10. Then, the experimenter tested if the child could repeat three-digit sequences in quiet using live speech. Next, the headphones were placed and a practice run was performed, starting at a favorable SNR (0 or −2 dB). The practice run was always performed in the N0S0$_{stat.}$ condition. The child repeated the triplets, and the experimenter recorded the response in the computer program. Next, up to eight SRTs were measured, by recording a test and a retest for the four different conditions. The order of testing the different conditions was counter balanced. First, a test and retest measurement was obtained for one of the conditions. Then, the other three conditions were tested. Finally, the retest measurements for these three conditions were obtained. The starting SNR for each test was chosen 6 to 8 dB more favorable than the expected SRT, and therefore depended slightly on the test condition and age of the child. On average, the starting level was −2 dB SNR for the N0S0$_{stat.}$ condition, −5 dB SNR for the N0S0$_{fluct.}$ and N0Sπ$_{stat.}$ conditions, and −7 dB SNR for the N0Sπ$_{fluct.}$ condition. The effect of the starting SNR on the SRT is very small or negligible, depending on the difference between SNR and the SRT (Smits & Houtgast 2006). A pure-tone audiogram was taken halfway during the testing session, and a short break was held after two or three tests. Testing sessions took 30 to 40 min. The session was ended if all tests were completed, or if the child indicated he/she wanted to stop or appeared to have lost attention. The experimenter motivated the younger children with verbal encouragement, and they received a sticker afterward. Testing sessions for adults were similar, except that they typed their response in the computer program themselves, and the pure-tone audiogram was determined before the DIN tests.

Results

Most children could perform DIN test and retest for multiple listening conditions after the practice run (on average five

<zdoi; 10.1097/AUD.0000000000000569>

TABLE 2. Group mean SRT scores and measurement error for the DIN test in various listening conditions for different age groups

Condition	Age Group	N	DIN SRT* (dB SNR)	DIN SEM (dB)
N0S0$_{stat.}$	4–6 yrs	33	−7.9±1.1	0.7
	7–9 yrs	49	−8.8±1.1	0.9
	10–12 yrs	25	−9.7±0.8	0.8
	Adult	33	−9.9±0.8	0.6
N0S0$_{fluct.}$	4–6 yrs	30	−11.5±1.5	0.9
	7–9 yrs	47	−12.7±1.5	1.0
	10–12 yrs	25	−13.8±1.0	1.0
	Adult	33	−15.3±1.1	1.1
N0Sπ$_{stat.}$	4–6 yrs	28	−10.9±1.5	0.9
	7–9 yrs	46	−12.8±1.4	1.1
	10–12 yrs	25	−14.2±1.2	0.9
	Adult	33	−15.1±0.9	0.8
N0Sπ$_{fluct.}$	4–6 yrs	30	−14.6±1.5	1.1
	7–9 yrs	47	−16.3±1.8	1.1
	10–12 yrs	25	−18.1±1.0	1.0
	Adult	33	−20.4±0.9	1.0

*Mean ± SD.
DIN, digits-in-noise; N0S0$_{fluct.}$, diotic presentation using fluctuating noise; N0S0$_{stat.}$, diotic presentation using stationary noise; N0Sπ$_{fluct.}$, dichotic presentation using fluctuating noise; N0Sπ$_{stat.}$, dichotic presentation using stationary noise; SEM, standard error of the mean; SNR, signal to noise ratio; SRT, speech reception threshold.

conditions for children <8 years of age, on average 7 conditions for children ≥8 years of age; adults completed all eight test and retest conditions). All except 1 of the 112 children (98%) completed the DIN test and retest for multiple conditions confirming that even young children (4 to 6 years of age) can do the DIN test without difficulty.

DIN Test Measurement Error for Children and Adults • Mean SRT scores and SEM for the DIN test are summarized in Table 2, for different age groups and in various listening conditions. Mean SRT scores were calculated by averaging test scores. The SEM was calculated from the test and retest when the retest results were available. Results of paired sample t tests (Bonferroni adjusted) for each age group (4–6 years, 7–9 years, 10–12 years, adults) indicated no statistically significant learning effect for any condition and any age group after the practice run, except for adults in the N0S0$_{fluct.}$ condition [0.7 dB learning effect; $t(32) = 2.529$; $p = 0.017$] and in the N0Sπ$_{stat.}$ condition [0.5 dB learning effect; $t(32) = 2.624$; $p = 0.013$]. The test–retest reliability expressed as SEM was approximately 1 dB for all conditions and age groups.

Age Dependency of the SRT • A one-way ANOVA was conducted to compare the effect of age group on SRT for the different conditions. There was a significant effect of age group on SRT for each condition tested [N0S0$_{stat.}$: $F(3,136) = 26.6$, $p < 0.001$; N0S0$_{fluct.}$: $F(3,131) = 46.4$, $p < 0.001$; N0Sπ$_{stat.}$: $F(3,128) = 59.5$, $p < 0.001$; N0Sπ$_{fluct.}$: $F(3,131) = 100.4$, $p < 0.001$] The SRT clearly depends on age for each condition tested, and younger children need a more favorable SNR to obtain similar speech recognition scores as adults.

Post hoc multiple comparisons with Bonferroni correction showed significant differences ($p < 0.01$) between all age groups for each condition, except for the difference in SRT between adults and the children 10 to 12 years of age in the N0S0$_{stat.}$ (mean difference, −0.22 dB; $p = 1.00$) and N0Sπ$_{stat.}$

(mean difference, −0.8 dB; $p = 0.1$) conditions. This means that children from roughly 10 years of age onward perform on the same level as adults in the stationary noise conditions. Incidentally, even individual children from the age of 6 performed similar to adults in these conditions. Only for the N0Sπ$_{fluct.}$ condition, the average of the adult SRTs is consistently better than any of the children's test score. For the N0S0$_{stat.}$, N0S0$_{fluct.}$, and N0Sπ$_{stat.}$ conditions, adults had a 4 to 5 dB more favorable SRT than 4-year-old children, while for the N0Sπ$_{fluct.}$ condition, the difference was 8 dB.

Figure 3 shows the SRT as a function of age for each condition tested. The individual test scores for each child are shown in a scatter plot. The retest score is not shown, so that each data point represents the performance of a single child with a single DIN test. Again the box plot represents the results for the adult group, with median group SRT (horizontal line), 25th and 75th percentile SRT (box ends), 10th and 90th percentile SRT (whiskers), and outliers (open circles). The thick lines are exponential fits to the data from the children. The regression equations are shown in Figure 3.

The intersubject variability in SRT values varied between conditions, both for children and adults. After correction for the age-dependent group mean SRT, the standard deviations for the N0S0$_{stat.}$, N0S0$_{fluct.}$, N0Sπ$_{stat.}$, and N0Sπ$_{fluct.}$ conditions were 1.0, 1.3, 1.3, and 1.4 dB for children, and 0.8, 1.3, 1.0, and 0.7 dB for adults, respectively. We calculated correlation coefficients across conditions in the residuals that remain after correction for the age-dependent mean SRT. For adults, we found no significant correlation across conditions. For children, we found significant ($p < 0.01$) moderate (r between 0.25 and 0.35) correlation coefficients across all conditions, except between the N0Sπ$_{fluct.}$ and N0S0$_{stat.}$ conditions ($r = 0.06$; $p = 0.27$). This means that roughly 10% (0.1 dB) of the variance can be explained by correlation between a child's test results across conditions. This is not the case for adults. Given the SEM of approximately 1 dB, the spread around the regression line can therefore largely be attributed to the measurement error.

Effect of Baseline SNR on Masking Release • We explored whether the age effects in children on the SRTs for the listening conditions with interrupted noise and N0Sπ presentation can be explained by their less favorable SRT in the N0S0$_{stat.}$ condition (baseline condition; Bernstein & Grant 2009). As explained in the introduction, this dependency occurs when the slopes of the speech recognition functions differ. Speech recognition functions for adults were constructed for the four conditions from the raw data (i.e., individual data points for each trial). Figure 4A shows the mean percentage correct versus SNR for each condition tested. The data show a characteristic sigmoidal pattern for each condition tested. The thick lines represent maximum likelihood fits of a logistic function. The slope of the speech recognition function at 50% correct is 14.5%/dB for N0S0$_{stat.}$; 10.6%/dB for N0S0$_{fluct.}$; 13.2%/dB for N0Sπ$_{stat.}$; and 12.8%/dB for N0Sπ$_{fluct.}$. The slope of the N0S0$_{fluct.}$ speech recognition function is significantly different from the slope of the baseline N0S0$_{stat.}$ speech recognition function (Wald $\chi^2 = 5.22$; $p < 0.05$). No significant difference between the slope of the N0S0$_{stat.}$ speech recognition function and the slopes of the N0Sπ$_{stat.}$ (Wald $\chi^2 = 0.51$; $p = 0.48$) and N0Sπ$_{fluct.}$ (Wald $\chi^2 = 0.99$; $p = 0.32$) speech recognition function was observed.

Figure 4B shows how the masking release for FMB, BU, and FMB&BU depends on the baseline SNR for adult subjects.
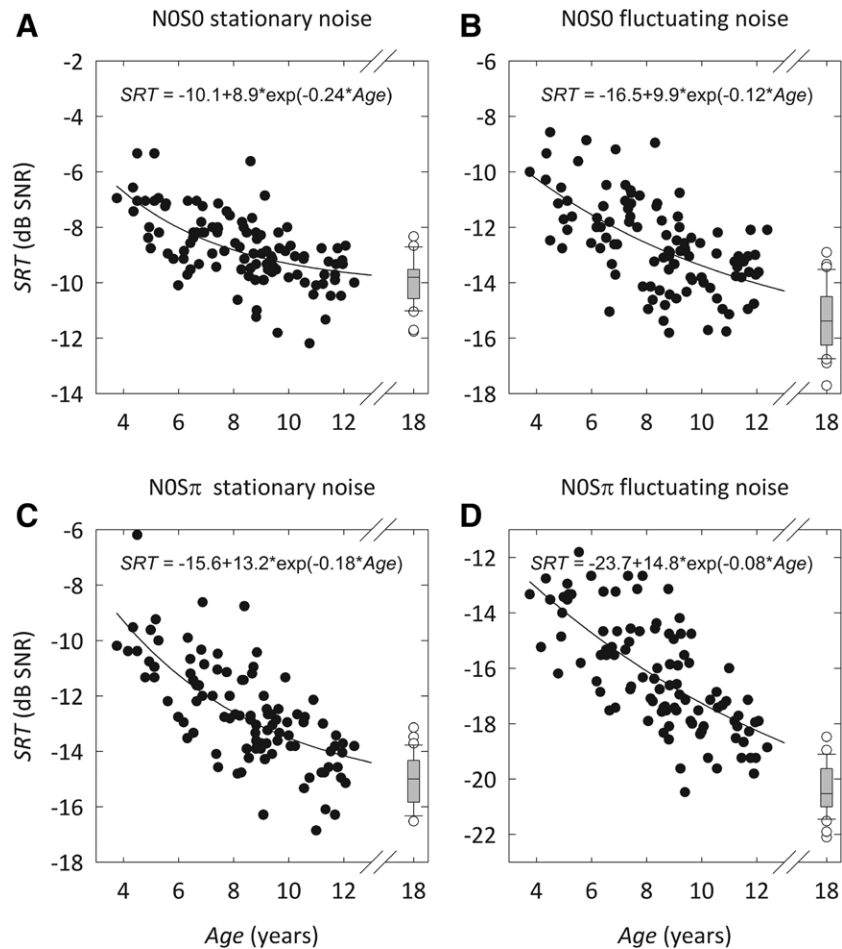
<zdoi; 10.1097/AUD.0000000000000569>

Fig. 3. DIN SRT as a function of age for different conditions: (A) N0S0 stationary noise, (B) N0S0 fluctuating noise, (C) N0Sπ stationary noise, (D) N0Sπ fluctuating noise. The thick lines are exponential fits to the data. The results from the first test for each condition are shown. The adult data are summarized in a box-whisker plot [median group SRT (horizontal line), 25th and 75th percentile SRT (box ends), 10th and 90th percentile SRT (whiskers), and outliers (open circles)]. DIN indicates digits-in-noise; N0S0, diotic; N0Sπ, dichotic; SNR, signal to noise ratio; SRT, speech reception threshold.

Because the slope for the baseline speech recognition function (N0S0$_{stat.}$) is steeper than for the N0S0$_{fluct.}$ speech recognition function, the amount of FMB is dependent on the baseline SNR at which the masking release is estimated (Bernstein & Grant 2009; Smits & Festen 2011, 2013). When, for instance, masking release is calculated relative to the baseline performance at −10 dB SNR (the observed SRT in the N0S0$_{stat.}$ condition for adults), the FMB is 5.7 dB. If, however, masking release is calculated relative to a baseline performance of −6 dB (a typical SRT in the N0S0$_{stat.}$ condition for the youngest children), the predicted FMB is only 4.4 dB (illustrated by the gray arrows in Fig. 4B). By contrast, the slopes for the N0Sπ$_{fluct.}$ and N0Sπ$_{stat.}$ speech recognition functions are not significantly different from the slope of the baseline speech recognition function N0S0$_{stat.}$. Therefore, BU and FMB&BU for adults do not significantly depend on the baseline SNR at which masking release is estimated, while FMB does depend on baseline SNR.

Age Effects of BU and FMB • The FMB measured in children should be compared with the baseline-corrected masking release for adults to detect possible true differences in masking release between children and adults. Baseline-corrected masking release for adults is calculated relative to a baseline SNR equal to the SRT of children in the various age groups. In other words, we compare the masking release experienced by a child, with the predicted masking release for adults at equal baseline SNR. Then, the difference in masking release can be attributed to an age effect.

The baseline-corrected masking release for adults is predicted from Figure 4B: (1) the N0S0$_{stat.}$ SRT for each child is taken as baseline SNR; (2) the baseline-corrected masking release is estimated from the difference in SNR between the baseline adult speech recognition function and other adult speech recognition function at the same percentage correct; and (3) the estimates for baseline-corrected masking release are averaged for each age group. The difference between the measured masking release in children and the baseline-corrected masking release for adults represents the true age-related difference in masking release between children and adults for each age group.

Figure 5 and Table 3 show the masking release for different age groups for all conditions (FMB, BU, and FMB&BU). Both the measured masking release (black circles) and the baseline-corrected masking release for adults (gray triangles) are shown. Our results show a release from masking for fluctuating noise versus stationary noise (FMB; Fig. 5A, black line) and for N0Sπ presentation versus N0S0 presentation (BU; Fig. 5B, black line). The combined effect of fluctuating noise and N0Sπ presentation versus stationary noise and N0S0 presentation even shows
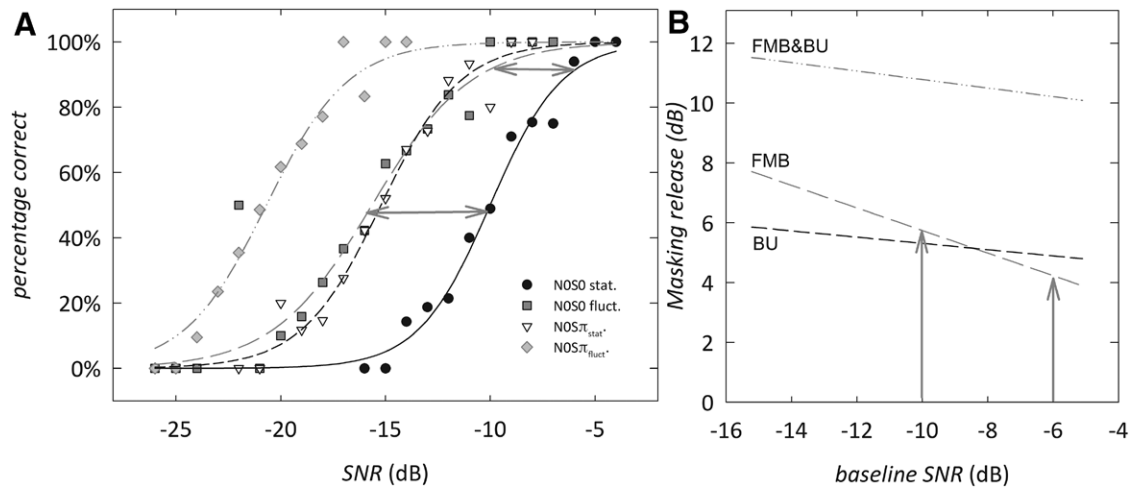
<zdoi; 10.1097/AUD.0000000000000569>

Fig. 4. Adult speech recognition functions and baseline-predicted masking release. A, Diotic and dichotic speech recognition functions for adult subjects for the four listening conditions tested. Data points show the mean percentage correct at each SNR. The lines show a logistic function fit to the individual data points. B, Masking release as a function of the baseline SNR for FMB, BU, and the combined effect. Note that there is a reduced release of masking at high SNR relative to low SNR, as illustrated with gray arrows for the FMB case. BU indicates binaural unmasking; DIN, digits-in-noise; FMB, fluctuating masker benefit; FMB&BU, the combined effect of fluctuating masker benefit and binaural unmasking; N0S0$_{fluct.}$, diotic presentation using fluctuating noise; N0S0$_{stat.}$, diotic presentation using stationary noise; N0Sπ$_{fluct.}$, dichotic presentation using fluctuating noise; N0Sπ$_{stat.}$, dichotic presentation using stationary noise; SNR, signal to noise ratio; SRT, speech reception threshold.

a greater release from masking (FMB&BU; Fig. 5C, black line). A one-way ANOVA was conducted to compare the effect of age group on FMB, BU, and FMB&BU. There was a significant effect of age on FMB [$F(3,131) = 10.5$; $p < 0.001$], BU [$F(3,125) = 12.1$; $p < 0.001$], and FMB&BU [$F(3,127) = 32.0$; $p < 0.001$]. A t test of measured masking release for children in different age groups and baseline-corrected masking release for adults showed a significant difference between measured masking release for all age groups and all types of masking release (FMB, BU, and FMB&BU), and baseline-corrected masking release for adults ($p < 0.05$). The only exception was BU for children 10 to 12 years of age, where no significant effect was observed [$t(23) = −1.93$; $p = 0.07$]. Only for fluctuating noise,

part of the reduced masking release observed with children can be explained by the lower SNR in the N0S0$_{stat.}$ condition for children. For each condition tested, adults have more masking release than children, even when accounting for baseline performance, as shown by the black line below the gray line in Figure 5. The only exception is BU for children 10 to 12 years of age, where adult-like performance is observed.

Discussion

Assessment of Speech Recognition in Noise Abilities in Children: pDIN and DIN Test Feasibility • The DIN test was designed to only minimally depend on nonauditory factors. It
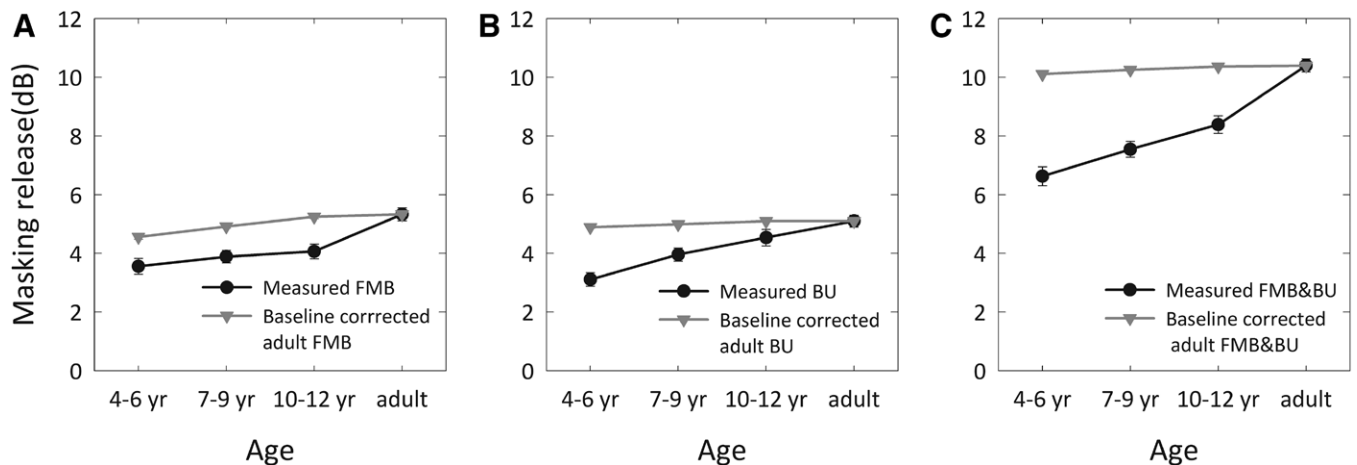

Fig. 5. Masking release (mean and standard error of the mean) for different age groups. The measured masking release (black line) is compared with the baseline-corrected masking release for adults (gray line), that is, the predicted masking release, an adult would experience relative to a baseline SNR equal to the SRT-DIN of the children in each age group. Results are shown for (A) FMB, (B) BU, and (C) FMB&BU. Only for fluctuating noise, part of the reduced masking release observed with children can be explained by the lower SNR in the N0S0$_{stat.}$ condition for children. For each condition tested (except BU in 10- to 12-year-olds), adults have more masking release than children, even when accounting for baseline performance. The error bars are smaller than the associated symbols in some cases. BU indicates binaural unmasking; DIN, digits-in-noise; FMB, fluctuating masker benefit; FMB&BU, the combined effect of fluctuating masker benefit and binaural unmasking; N0S0$_{stat.}$, diotic presentation using stationary noise; SNR, signal to noise ratio; SRT, speech reception threshold.

<zdoi; 10.1097/AUD.0000000000000569>

TABLE 3. Group mean measured masking release (FMB, BU, and FMB&BU) for different age groups compared with baseline corrected masking release for adults

Age Group	Type of Masking Release	Measured Masking Release (dB)*	Baseline-Corrected Masking Release for Adults (dB)*	Mean Difference (dB)
4–6 yrs	FMB	3.6 ± 1.5	4.6 ± 0.4	−1.0†
	BU	3.1 ± 1.2	4.9 ± 0.13	−1.8†
	FMB&BU	6.7 ± 1.6	10.1 ± 0.16	−3.5†
7–9 yrs	FMB	3.9 ± 1.5	4.9 ± 0.4	−1.0†
	BU	4.0 ± 1.5	5.0 ± 0.12	−1.0†
	FMB&BU	7.6 ± 1.8	10.3 ± 0.16	−2.7†
10–12 yrs	FMB	4.1 ± 1.2	5.2 ± 0.3	−1.2†
	BU	4.5 ± 1.4	5.1 ± 0.09	−0.5
	FMB&BU	8.4 ± 1.5	10.4 ± 0.12	−2.0†
Adult	FMB	5.3 ± 1.3	—	—
	BU	5.1 ± 1.1	—	—
	FMB&BU	10.4 ± 1.3	—	—

*Mean ± SD.
†Significant at the $p < 0.005$ level.
BU, binaural unmasking; FMB, fluctuating masker benefit; FMB&BU, the combined effect of fluctuating masker benefit and binaural unmasking.

could however be challenging for young children. Therefore, this study compared the DIN test with a pDIN test that aims to rule out potential remaining nonauditory factors in the test that might influence test performance by the youngest children. These factors include difficulty with the abstract digit 0 and the auditory memory needed for repeating three-digit sequences. Our results demonstrate that even young children from 4 years of age could perform both pDIN test and DIN test. All children were familiar with the speech material (numbers 0 to 9), and we did not find evidence for extra difficulty with the digit 0. Most children could perform the task of repeating digits or triplets and could complete all the tests. After a practice run, there was no significant learning effect for children and adults. The measurement error of the DIN test was smaller than that of the pDIN test, and comparable to the measurement error found by Smits et al. (2013) for adults. The smaller measurement error relates to the steeper speech recognition function for triplets relative to digits (Smits et al. 2013; Smits & Houtgast 2006), and to the very low guess rate on the DIN test.

A small fraction of the children and adults reported hearing four digits instead of three in the DIN test. Typically, an extra 0 or "1" was heard before the triplet. This was probably due to an auditory illusion of an extra digit in the noise onset. Careful reinstruction to ignore the extra digit helped to complete the test in all adults and most of the children. Only 2 children were too confused to do the test because of this phenomenon and were not included in this study. This auditory illusion was not present in case of the pDIN test, where only a single digit was presented.

Normative data from the digit span tests in the Wechsler Intelligence Scale for Children show that the digits span increases with age and that 1.5% of the 6-year-old children cannot recall three digits in the forward direction (Wechsler 2004). This suggests that a small fraction of the 6-year-old and probably even larger fractions of 4- and 5-year-old children do not have the auditory memory to do the DIN test reliably. However, our results did not demonstrate this effect in our test population

which may be due to the small sample size of 4- and 5- year-olds in our experiment. We found no significant difference between pDIN SRT and DIN SRT for adults and children, and there was a strong correlation between both SRTs. Thus, despite the differences between DIN test and pDIN test in test procedure, task, target point, and speech recognition function, the results suggest that both tests essentially measure the same auditory ability and the test results can be used interchangeably.

It can be concluded that the use of the DIN test is feasible for normal-hearing children from 4 years of age. A pediatric version of the test with specific modifications to simplify the test such as the use of single digits instead of triplets, the omission of the digit 0, and animations to promote the attention span for young children are not necessary. The DIN test is preferable over the pDIN test because of the smaller measurement error and small intersubject variability. Further research is needed to explore whether the DIN test can also be used for hearing-impaired children from that age onward, and for even younger children, who may have a reduced attention span and auditory memory.

SRT Dependency on Age: Auditory and Nonauditory Factors • For all tests and test conditions, we found a significant improvement of the SRT with age. Adult-like performance is reached at the age of 10 to 12 years in the stationary noise conditions. These findings are consistent with those reported in the literature (Garadat & Litovsky 2007; Hall et al. 2002; Holder et al. 2016; Nishi et al. 2010; Stuart 2008; Vaillancourt et al. 2008; Van Deun et al. 2010; Yuen & Yuan 2014), showing an improvement in SRT for children with age, and achieving adult-like performance in stationary noise conditions by the age of at least 10 years. The origin of this age-related improvement of the SRT is not completely understood. It has been a subject of debate to what extent various auditory and nonauditory factors play a role (Moore et al. 2011). It has been argued that the observed improvement of SRTs is accompanied with a maturation of auditory pathways and binaural processing (e.g., Eggermont & Ponton 2003), and therefore reflects an ongoing maturation of auditory perception. However, the development of linguistic and cognitive skills, working memory, and attention skills takes place simultaneously and may play a role in assessing speech recognition abilities (Elliott 1979). Several studies demonstrate that children's ability to recognize speech appears to take longer to mature and follows a different developmental trajectory for two-talker speech maskers than for speech-shaped noise maskers (Buss et al. 2016; Corbin et al. 2016; Leibold & Buss 2013). Perceptual maskers (competing speech) are thought to place a greater processing load on both the auditory and cognitive systems and rely on executive function such as attention and working memory (Corbin et al. 2016). Jones et al. (2015) evaluated whether changes in internal noise or in selective attention could explain the development of hearing in noise in children. They found that the improvement in selective attention alone was the only important mechanism. This suggests that observed developmental changes in hearing in noise for children 4 years of age and onward are most likely nonsensory in origin.

In the present study, however, we aimed to minimize the dependency on these nonauditory factors by using the DIN test. The test was designed to minimize the influence of nonauditory factors on the SRT (Smits et al. 2013). Kaandorp et al. (2016) showed that the stimuli used in the DIN test only require a minimum of linguistic skills. Heinrich et al. (2015)

<zdoi; 10.1097/AUD.0000000000000569>

reported that digit–triplet recognition by adults was influenced by hearing sensitivity alone, and not by cognition. This agrees with the findings of Talarico et al. (2007), who found that children with higher cognitive abilities did not outperform children with lower cognitive abilities on speech-in-noise tasks involving speech-shaped noise. By contrast, Moore et al. (2014) did find a close association between cognition and DIN SRT for older adults. Their study does not report the use of a practice run before the test, however. Smits et al. (2013) showed that a single practice list is necessary for naive listeners to eliminate procedural learning effects. The effect size of this initial learning could depend on the subject's cognition, and we speculate that this could play a role in the findings by Moore et al. The present study shows that the speech material in the DIN test is completely familiar even to the youngest children. Furthermore, the task is simple and straightforward. Most children 6 years of age and onward have the necessary auditory memory skills for a digit span of three digits, which is required to perform the DIN test (Wechsler 2004). Our results suggest that auditory memory is not a limiting factor in DIN test performance, as shown by the similarity of SRTs in the DIN test and pDIN test.

In summary, it is unlikely that the observed improvement of DIN SRT with age in children can be explained by the ongoing development of cognitive skills, linguistic skills, and auditory memory alone. It is likely that both auditory factors, such as the ongoing maturation of the auditory system, and nonauditory factors, such as the development of selective attention, play a role.

FMB and BU • The present study shows that auditory abilities required to benefit from fluctuating maskers and binaural cues develop well into adolescence. FMB has been studied extensively with adult subjects (Festen & Plomp 1990; Rhebergen et al. 2006), but the number of studies with children is limited. Stuart (2005) found an age-dependent FMB when testing speech recognition in noise with words in children and attributed this to the improving temporal resolution by ongoing maturation of central auditory processing. However, FMB did not improve with age when testing speech recognition in noise with sentences in children (Stuart 2008), and he concluded that school-age children have inherently poorer processing efficiency. Because of the nature of the speech material in these tests (words and sentences, respectively), a linguistic component in these findings cannot be ruled out. Also, baseline performance was not considered in these studies. Studies by Hall et al. (2012) and Buss et al. (2016) showed a significant developmental effect on masking release related to the temporal modulations in the noise. They speculated that the observed age effect might be related to the difference in baseline performance between children and adults.

Figure 5 shows an increase of FMB with age, even when correcting for the difference in baseline performance between children and adults. It is not clear whether this increase in FMB can be attributed to an improving temporal resolution by ongoing maturation of central auditory processing (Stuart 2005), or to other causes such as development in selective attention (Jones et al. 2015) for speech fragments in the dips of the fluctuating noise. Nevertheless, young children seem to be relatively poor at piecing together sparse glimpses of speech.

Figure 5 also demonstrates that BU improves with age, and adult-like performance is not achieved before the age of 10 years. BU has been studied with adults before (George et al. 2012; George et al. 2010; Goverts & Houtgast 2010; Johansson & Arlinger 2002), but binaural processing with children is mostly investigated with a spatial release from masking paradigm. SRM is calculated by the release from masking when speech is spatially separated from noise in a free-field testing setup. Unlike BU, which only depends on a binaural phase difference, SRM also depends on head shadow effects and interaural level differences. As discussed in the introduction, some studies report that SRM does not improve with age (Ching et al. 2011; Litovsky 2005; Murphy et al. 2011) while other studies show a significant improvement with age (Vaillancourt et al. 2008; Van Deun et al. 2010; Yuen & Yuan 2014). In a different testing paradigm, where speech and noise are simulated to be spatially separated under a headphone, the SRM also significantly improves with age (Cameron et al. 2009; Cameron & Dillon 2007). Yuen & Yuan (2014) concluded that it is likely that the observed developmental time course of BU is related to ongoing development of binaural auditory processing at the auditory brainstem (Moore, 1991) or even cortical levels. Our results are in agreement with the latter studies, and show an age-related improvement in the ability to benefit from binaural cues in children.

The present study supports the notion that young children need a favorable SNR in acoustically demanding listening situations, such as at school: their ability to separate speech from noise is still developing, as is their ability to benefit from masker fluctuations and binaural cues. The age at which children achieve adult-like performance is different for the conditions tested. This may reflect the fact that, although perceptual maturation generally seems to extend well into adolescence, the age at which adult-like performance is attained may be different for different auditory abilities (Sanes & Woolley 2011). For almost all tests and test conditions in the present experiment, adults have on average better SRT scores and release from masking than children.

On top of these developmental effects on auditory abilities, children with hearing loss suffer from a compromised auditory system, resulting in poorer speech recognition abilities and more listening effort in adverse listening conditions (Hick & Tharpe 2002). Therefore, children with hearing loss need even a more favorable SNR than their normal-hearing peers. These data highlight the need for ensuring good classroom acoustics for all children in primary schools.

Clinical Application of the DIN Test in Children • The application of the DIN test with children in clinical practice is feasible because most children were able to carry out the test and the testing time was only 3 to 4 min. By measuring the age-dependent SRT for a large group (N = 112) of normal-hearing children between 4 and 12 years of age, we could establish normative age-dependent data (mean and 95% percentile) for headphone testing, making the test ready for clinical use for monotic, N0S0, and N0Sπ presentation, and in stationary and fluctuating noise. The DIN test can be adapted to incorporate other relevant conditions (e.g., reverberation) or presentation modes (e.g., free-field presentation, different headphones) but then correction factors to the normative data have to be established.

It has been demonstrated that the DIN test is applicable to adults and the elderly population (Smits et al. 2013), cochlear implant recipients and hearing aid users (Kaandorp et al. 2015), nonnative listeners (Kaandorp et al. 2016), and normal-hearing children (present study). Therefore, speech recognition in noise

abilities can be tested with a single test that is suitable for most clinical populations. Further research is needed to understand how hearing loss or language impairment in children affects test performance on the DIN test. Finally, this test could then potentially be used as a simple and elegant instrument for hearing screening at primary schools.

CONCLUSIONS

Speech recognition abilities in children in acoustically demanding listening conditions can accurately and reliably be tested using the DIN test. A pediatric single-digit version of the test is not necessary for children over 4 years of age, making the DIN test applicable to a wide clinical population. Speech recognition in noise abilities develop well into adolescence, and young children need a more favorable SNR than adults for all listening conditions tested (stationary noise, interrupted noise, monotic, N0S0, and N0Sπ presentation). Older children have SRTs comparable to those of adults in stationary noise conditions. It is unlikely that the observed improvement of DIN SRT with age in children can be explained by the ongoing development nonauditory factors alone. Children gain less benefit from fluctuating maskers and from binaural cues than adults, even when corrected for adult baseline performance. We established age-dependent normative data for the Dutch version of the DIN test for various listening conditions, based on data of more than 100 children 4 to 12 years of age.

REFERENCES

Bernstein, J. G., & Grant, K. W. (2009). Auditory and auditory-visual intelligibility of speech in fluctuating maskers for normal-hearing and hearing-impaired listeners. *J Acoust Soc Am*, *125*, 3358–3372.

Buss, E., Hall, J. W. 3rd, Grose, J. H. (2006). Development and the role of internal noise in detection and discrimination thresholds with narrow band stimuli. *J Acoust Soc Am*, *120*(5 Pt 1), 2777–2788.

Buss, E., Leibold, L. J., Hall, J. W. 3rd. (2016). Effect of response context and masker type on word recognition in school-age children and adults. *J Acoust Soc Am*, *140*, 968–977.

Cameron, S., & Dillon, H. (2007). Development of the listening in spatialized noise-sentences test (LISN-S). *Ear Hear*, *28*, 196–211.

Cameron, S., Brown, D., Keith, R., et al. (2009). Development of the North American Listening in Spatialized Noise-Sentences test (NA LiSN-S): Sentence equivalence, normative data, and test-retest reliability studies. *J Am Acad Audiol*, *20*, 128–146.

Ching, T. Y., van Wanrooy, E., Dillon, H., et al. (2011). Spatial release from masking in normal-hearing children and children who use hearing aids. *J Acoust Soc Am*, *129*, 368–375.

Ching, T. Y., Zhang, V. W., Flynn, C., et al. (2017). Factors influencing speech perception in noise for 5-year-old children using hearing aids or cochlear implants. *Int J Audiol*. Jul. 7. [Epub ahead of print].

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

Crandell, C. C. (1993). Speech recognition in noise by children with minimal degrees of sensorineural hearing loss. *Ear Hear*, *14*, 210–216.

Eggermont, J. J., & Ponton, C. W. (2003). Auditory-evoked potential studies of cortical maturation in normal hearing and implanted children: Correlations with changes in structure and speech perception. *Acta Otolaryngol*, *123*, 249–252.

Elliott, L. L. (1979). Performance of children aged 9 to 17 years on a test of speech intelligibility in noise using sentence material with controlled word predictability. *J Acoust Soc Am*, *66*, 651–653.

Festen, J. M., & Plomp, R. (1990). Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing. *J Acoust Soc Am*, *88*, 1725–1736.

Garadat, S. N., & Litovsky, R. Y. (2007). Speech intelligibility in free field: Spatial unmasking in preschool children. *J Acoust Soc Am*, *121*, 1047–1055.

George, E. L., Festen, J. M., Goverts, S. T. (2012). Effects of reverberation and masker fluctuations on binaural unmasking of speech. *J Acoust Soc Am*, *132*, 1581–1591.

George, E. L., Goverts, S. T., Festen, J. M., et al. (2010). Measuring the effects of reverberation and noise on sentence intelligibility for hearing-impaired listeners. *J Speech Lang Hear Res*, *53*, 1429–1439.

Goverts, S. T., & Houtgast, T. (2010). The binaural intelligibility level difference in hearing-impaired listeners: The role of supra-threshold deficits. *J Acoust Soc Am*, *127*, 3073–3084.

Hall, J. W., Buss, E., Grose, J. H., et al. (2004). Developmental effects in the masking-level difference. *J Speech Lang Hear Res*, *47*, 13–20.

Hall, J. W., Buss, E., Grose, J. H., et al. (2012). Effects of age and hearing impairment on the ability to benefit from temporal and spectral modulation. *Ear Hear*, *33*, 340–348.

Hall, J. W. 3rd, Grose, J. H., Buss, E., et al. (2002). Spondee recognition in a two-talker masker and a speech-shaped noise masker in adults and children. *Ear Hear*, *23*, 159–165.

Heinrich, A., Henshaw, H., Ferguson, M. A. (2015). The relationship of speech intelligibility with hearing sensitivity, cognition, and perceived hearing difficulties varies for different speech perception tests. *Front Psychol*, *6*, 782.

Hick, C. B., & Tharpe, A. M. (2002). Listening effort and fatigue in school-age children with and without hearing loss. *J Speech Lang Hear Res*, *45*, 573–584.

Holder, J. T., Sheffield, S. W., Gifford, R. H. (2016). Speech understanding in children with normal hearing: Sound field normative data for Baby-Bio, BKB-SIN, and QuickSIN. *Otol Neurotol*, *37*, e50–e55.

Johansson, M. S., & Arlinger, S. D. (2002). Binaural masking level difference for speech signals in noise. *Int J Audiol*, *41*, 279–284.

Jones, P. R., Moore, D. R., Amitay, S. (2015). Development of auditory selective attention: Why children struggle to hear in noisy environments. *Dev Psychol*, *51*, 353–369.

Kaandorp, M. W., De Groot, A. M., Festen, J. M., et al. (2016). The influence of lexical-access ability and vocabulary knowledge on measures of speech recognition in noise. *Int J Audiol*, *55*, 157–167.

Kaandorp, M. W., Smits, C., Merkus, P., et al. (2015). Assessing speech recognition abilities with digits in noise in cochlear implant and hearing aid users. *Int J Audiol*, *54*, 48–57.

Kaernbach, C. (1991). Simple adaptive testing with the weighted up-down method. *Percept Psychophys*, *49*, 227–229.

Leibold, L. J., & Buss, E. (2013). Children's identification of consonants in a speech-shaped noise or a two-talker masker. *J Speech Lang Hear Res*, *56*, 1144–1155.

Licklider, J. C. R. (1948). The influence of interaural phase relations upon the masking of speech by white noise. *J Acoust Soc Am*, *20*, 150–159.

Litovsky, R. Y. (2005). Speech intelligibility and spatial release from masking in young children. *J Acoust Soc Am*, *117*, 3091–3099.

<zdoi; 10.1097/AUD.0000000000000569>

Mendel, L. L. (2008). Current considerations in pediatric speech audiometry. *Int J Audiol*, *47*, 546–553.

Moore, D. R. (1991). Anatomy and physiology of binaural hearing. *Audiology*, *30*, 125–134.

Moore, D. R., Cowan, J. A., Riley, A., et al. (2011). Development of auditory processing in 6- to 11-yr-old children. *Ear Hear*, *32*, 269–285.

Moore, D. R., Edmondson-Jones, M., Dawes, P., et al. (2014). Relation between speech-in-noise threshold, hearing loss and cognition from 40-69 years of age. *PLoS One*, *9*, e107720.

Murphy, J., Summerfield, A. Q., O'Donoghue, G. M., et al. (2011). Spatial hearing of normally hearing and cochlear implanted children. *Int J Pediatr Otorhinolaryngol*, *75*, 489–494.

Neuman, A. C., Wroblewski, M., Hajicek, J., et al. (2010). Combined effects of noise and reverberation on speech recognition performance of normal-hearing children and adults. *Ear Hear*, *31*, 336–344.

Nishi, K., Lewis, D. E., Hoover, B. M., et al. (2010). Children's recognition of American English consonants in noise. *J Acoust Soc Am*, *127*, 3177–3188.

Oxenham, A. J., & Simonson, A. M. (2009). Masking release for low- and high-pass-filtered speech in the presence of noise and single-talker interference. *J Acoust Soc Am*, *125*, 457–468.

Plomp, R., & Mimpen, A. M. (1979). Improving the reliability of testing the speech reception threshold for sentences. *Audiology*, *18*, 43–52.

Rhebergen, K. S., Versfeld, N. J., Dreschler, W. A. (2006). Extended speech intelligibility index for the prediction of the speech reception threshold in fluctuating noise. *J Acoust Soc Am*, *120*, 3988–3997.

Sanes, D. H., & Woolley, S. M. (2011). A behavioral framework to guide research on central auditory development and plasticity. *Neuron*, *72*, 912–929.

Smits, C., & Festen, J. M. (2011). The interpretation of speech reception threshold data in normal-hearing and hearing-impaired listeners: Steady-state noise. *J Acoust Soc Am*, *130*, 2987–2998.

Smits, C., & Festen, J. M. (2013). The interpretation of speech reception threshold data in normal-hearing and hearing-impaired listeners: II. Fluctuating noise. *J Acoust Soc Am*, *133*, 3004–3015.

Smits, C., & Houtgast, T. (2006). Measurements and calculations on the simple up-down adaptive procedure for speech-in-noise tests. *J Acoust Soc Am*, *120*, 1608–1621.

Smits, C., Theo Goverts, S., Festen, J. M. (2013). The digits-in-noise test: Assessing auditory speech recognition abilities in noise. *J Acoust Soc Am*, *133*, 1693–1706.

Stuart, A. (2005). Development of auditory temporal resolution in school-age children revealed by word recognition in continuous and interrupted noise. *Ear Hear*, *26*, 78–88.

Stuart, A. (2008). Reception thresholds for sentences in quiet, continuous noise, and interrupted noise in school-age children. *J Am Acad Audiol*, *19*, 135–146; quiz 191.

Talarico, M., Abdilla, G., Aliferis, M., et al. (2007). Effect of age and cognition on childhood speech in noise perception abilities. *Audiol Neurootol*, *12*, 13–19.

Vaillancourt, V., Laroche, C., Giguère, C., et al. (2008). Establishment of age-specific normative data for the Canadian French version of the hearing in noise test for children. *Ear Hear*, *29*, 453–466.

Van Deun, L., van Wieringen, A., Wouters, J. (2010). Spatial speech perception benefits in young children with normal hearing and cochlear implants. *Ear Hear*, *31*, 702–713.

Versfeld, N. J., Daalder, L., Festen, J. M., et al. (2000). Method for the selection of sentence materials for efficient measurement of the speech reception threshold. *J. Acoust. Soc. Am, 107,* 1671–1684.

Wechsler, D. (2004). *The Wechsler intelligence scale for children* (4th ed.). London, United Kingdom: Pearson.

Wellman, H. M., Miller, K. F. (1986). Thinking about nothing: Development of concepts of zero. *Br J Dev Psychol*, *4*, 31–42.

Wilson, R. H., Farmer, N. M., Gandhi, A., et al. (2010). Normative data for the words-in-noise test for 6- to 12-year-old children. *J Speech Lang Hear Res*, *53*, 1111–1121.

Yuen, K. C., & Yuan, M. (2014). Development of spatial release from masking in mandarin-speaking children with normal hearing. *J Speech Lang Hear Res*, *57*, 2005–2023.