JASA ARTICLE

An analytical method to convert between speech recognition thresholds and percentage-correct scores for speech-in-noise tests

Cas Smits,[1,a)] Karina C. De Sousa,[2,b)] and De Wet Swanepoel[2,c)]
[1]Amsterdam UMC, Vrije Universiteit Amsterdam, Otolaryngology—Head and Neck Surgery, Ear & Hearing, Amsterdam Public Health Research Institute, De Boelelaan 1117, Amsterdam, The Netherlands
[2]Department of Speech-Language Pathology and Audiology, University of Pretoria, Pretoria, Gauteng, South Africa

ABSTRACT:
Speech-in-noise tests use fixed signal-to-noise ratio (SNR) procedures to measure the percentage of correctly recognized speech items at a fixed SNR or use adaptive procedures to measure the SNR corresponding to 50% correct (i.e., the speech recognition threshold, SRT). A direct comparison of these measures is not possible yet. The aim of the present study was to demonstrate that these measures can be converted when the speech-in-noise test meets specific criteria. Formulae to convert between SRT and percentage-correct were derived from basic concepts that underlie standard speech recognition models. Information about the audiogram is not being used in the proposed method. The method was validated by comparing the direct conversion by these formulae with the conversion using the more elaborate Speech Intelligibility Index model and a representative set of 60 audiograms ($r = 0.993$ and $r = 0.994$, respectively). Finally, the method was experimentally validated with the Afrikaans sentence-in-noise test ($r = 0.866$). The proposed formulae can be used when the speech-in-noise test uses steady-state masking noise that matches the spectrum of the speech. Because pure tone thresholds are not required for these calculations, the method is widely applicable. https://doi.org/10.1121/10.0005877

(Received 2 March 2021; revised 20 July 2021; accepted 22 July 2021; published online 24 August 2021)

[Editor: Karen S. Helfer] Pages: 1321–1331

I. INTRODUCTION

It has been recognized for decades that the negative effect of background noise on speech recognition is greater for listeners with hearing loss than for listeners with normal hearing (Carhart and Tillman, 1970; Plomp, 1978). Different approaches have been taken in the development of speech-in-noise tests to quantify speech recognition abilities in noise. Speech-in-noise-tests most often use either fixed-signal-to-noise ratio (SNR) procedures or adaptive procedures (Ricketts *et al.*, 2019).

Fixed-SNR (i.e., the constant-stimulus method) means that the SNR is fixed and all stimuli are presented at the same SNR. The test measures the percentage-correct at that specific SNR, thus, only one point on the speech recognition function is determined. The speech recognition function describes the relationship between the SNR and the percentage of correctly recognized speech items. The advantage of the procedure is the simplicity to administer the test and the result, a percentage-correct, which is an easily understood measure for researchers, clinicians, and patients. Examples of this type of speech-in-noise test are the BKB-SIN, Bamford-Kowal-Bench Speech-in-Noise test (Bench *et al.*, 1979), SPIN (Speech Perception in Noise test) (Bilger *et al.*, 1984), PRESTO, Perceptually Robust English Sentence Test Open-set (Gilbert *et al.*, 2013), and SPRINT (Speech Reception in Noise Test) (Brungart *et al.*, 2017). Disadvantages of the procedure are ceiling and floor effects for good and poor performers, respectively, and the difference score (e.g., comparing the effect of two different hearing aid settings or when comparing listeners) depending on the chosen SNR (Owen, 1981). Furthermore, with the fixed-SNR procedure, it is not possible to quantify differences in speech recognition abilities when measurements are performed at different SNRs.

Adaptive procedures determine the SNR where the average speech recognition score equals a certain percentage-correct, commonly called the target point. The one-up one-down adaptive procedure (Levitt, 1971; García-Pérez, 1998; Smits and Houtgast, 2006) is often used to measure the speech recognition threshold (SRT),[1] which represents the SNR where the listener recognizes 50% of the speech items correctly (sometimes denoted as SNR_{50}). Examples of adaptive speech-in-noise tests are the HINT (Hearing in Noise Test) (Nilsson *et al.*, 1994), the WIN (Words-in-Noise test) (Wilson, 2003), QuickSIN (Quick speech-in-noise test) (Killion *et al.*, 2004), DIN (Digits-in-Noise test) (Smits *et al.*, 2013), and MATRIX tests (Kollmeier *et al.*, 2015). The results of a speech-in-noise test can, therefore, be expressed as either percentage-correct (%), or, in dB SNR depending on the specific test and

a)Electronic mail: c.smits@amsterdamumc.nl, ORCID: 0000-0003-1617-1253.
b)ORCID: 0000-0003-1742-1613.
c)ORCID: 0000-0001-8313-1636.

measurement procedure. We are not aware of direct calculation methods to convert between percentage-correct and dB SNR, making direct comparison of results impossible.

The Speech Intelligibility Index (SII) model (ANSI, 1997) can be used to compute a physical measure (SII value), which represents the amount of speech information available to the listener depending on the speech signal, noise signal, and hearing thresholds of the listener. The method is relatively complex with numerous parameters and calculation steps, which makes it almost necessary to use computer software. The calculated SII value is related to the speech recognition score through a transfer function. When the SRT for a listener is known, it is possible to calculate the SII value using the hearing thresholds and properties of the speech and noise. Then the percentage-correct score at a certain SNR can be predicted using the SII model and the transfer function. Thus, converting SRTs to percentage-correct is possible, albeit complex, by using the SII model. Details about the SII model and the calculations are described below (Sec. II).

In the present study, an analytical method is presented to convert between results from a fixed-SNR procedure, expressed in percentage-correct and SRT. A unique aspect of the proposed method is that, unlike the SII method, pure tone thresholds are not required for the calculation. This makes it possible, for example, to quantify differences between groups of listeners or differences between conditions of published data in which fixed-SNR methods with different SNRs were used. The proposed conversion method is compared to the SII model, which uses pure tone thresholds to account for audibility, and a more elaborated SII model, which includes the effect of suprathreshold deficits. Further, an experimental study will be presented in which the conversion method is validated using a sentence-in-noise test presented to listeners with a wide range of hearing losses.

II. CONVERTING BETWEEN PERCENTAGE-CORRECT AND SRT I

Smits and Festen (2011, 2013) described speech recognition as a three-stage process. Their description largely followed the basic concepts underlying most of the present models of speech recognition: the articulation index (AI) (ANSI, 1986), the SII (ANSI, 1997), and the STI (Steeneken and Houtgast, 1980). Stage 1 concerns the physical description of the acoustical input signal. Stage 2 describes how well the peripheral auditory system can extract speech information from the input signal, and stage 3 describes the processing of the extracted speech information (i.e., top-down processing). For hearing-impaired listeners, the ability to extract speech information from the input signal is diminished, which is often interpreted as a problem in stage 2 of the recognition process. The impairments of the peripheral auditory system can be characterized either as a loss of audibility of a proportion of the input signal due to an elevated hearing threshold or limitation in extracting speech information from the audible portion of the input signal (i.e., suprathreshold deficits). As a consequence, the hearing-impaired listener needs an input signal with more speech information available (e.g., a broader bandwidth or a more favourable SNR) to achieve the same percentage-correct score as normal-hearing listeners. If the input signal for the normal-hearing listener and the hearing-impaired listener is identical, then percentage-correct scores will generally be lower for the hearing-impaired listener.

The SII model provides a standardized method of determining the intelligibility of speech by calculating the audibility in each frequency band and adding the weighted results. The audibility is expressed as the proportion of speech above the greater of either hearing threshold or the masking noise. The SII assumes that the audibility ranges from 0 to 1 when the SNR ranges from –15 to 15 dB, thus assuming a 30 dB dynamic range for speech. A level-distortion factor and self-speech masking are incorporated in the SII model. The level-distortion factor accounts for the decrease in intelligibility at high presentation levels, and self-speech masking refers to masking of higher speech frequencies by lower speech frequencies (ANSI, 1997). A desensitization factor has been proposed to correct for the suprathreshold deficits that hearing-impaired listeners may have (e.g., Pavlovic *et al.*, 1986; Ching *et al.*, 1998).

A transfer function denotes the relationship between SII values and percentage-correct score. The transfer function depends on the speaker and the kind of speech material used. The transfer function may also depend on the listener (non-native listeners, for instance, need more speech information than native listeners to achieve a certain percentage-correct score), but in standard SII calculations, one specific transfer function is used for each type of speech material. An SII value of 1 means that all the speech information is available to the listener, which corresponds to the upper asymptote of the transfer function (often 100%, but that can also be a lower score).

In the current paper, as in Smits and Festen (2011, 2013), an approximation of the SII model is used, which only holds for speech recognition in steady-state speech spectrum shaped noise. Two general properties of the SII function follow. First, the SII function for normal-hearing listeners maps SII values ranging from 0 to 1 linearly to the SNRs from –15 to 15 dB. Second, the SII function for hearing-impaired listeners is shallower and grows linearly from 0 to its maximum value, between SNR = –15 dB and SNR = 15 dB. See Smits and Festen (2011) for detailed information. Figure 1(A) shows the SII function for normal-hearing listeners and an example SII function for a hearing-impaired listener. The corresponding speech recognition functions are shown in Fig. 1(B). Note that hearing-impaired listeners may reach scores of 100% because, in general, SII values less than 1 are required to reach 100% recognition scores. Smits and Festen (2011) demonstrated that the hearing-impaired speech recognition function could be fully determined from the normal-hearing speech recognition function when the SRT for the hearing-impaired listener is known. They demonstrated this framework's

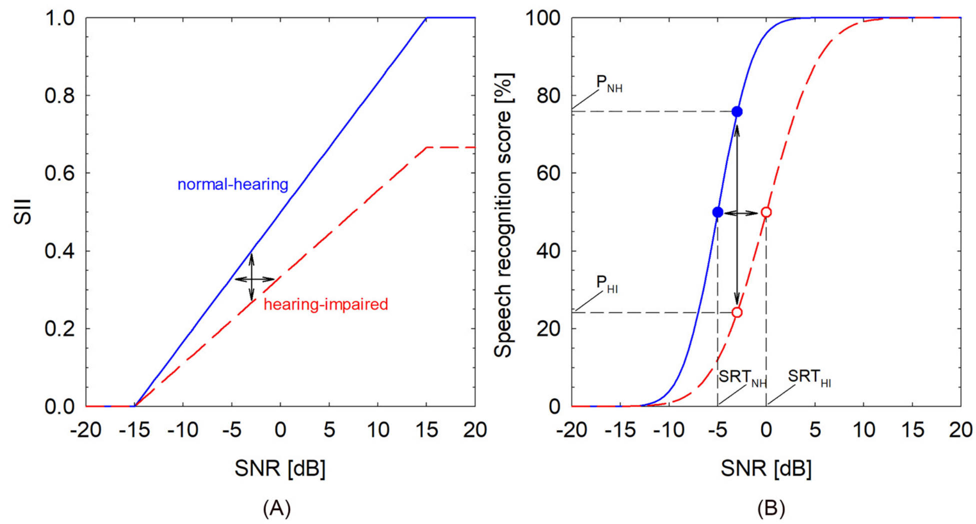

FIG. 1. (Color online) (A) The SII function for normal-hearing listeners and a sample SII function for a hearing-impaired listener in steady-state noise and (B) speech recognition functions for these listeners. The horizontal arrow shows the difference in SRT between normal-hearing listeners and a hearing-impaired listener. The vertical arrow represents the difference in percentage-correct at an SNR of −3 dB. The corresponding arrows are shown in (A).

validity by analyzing data sets from Smits and Houtgast (2006; nearly 40 000 digit-triplet SRTs from the Dutch National Hearing Test) and Zekveld *et al.* (2011; sentence SRTs). Because the entire hearing-impaired speech recognition function can be described analytically when only the SRT of the hearing-impaired listener and the normal-hearing speech recognition function are known, any conversion from SNR to percentage-correct can be done.

The arrows in Fig. 1(B) represent the results from two hypothetical speech-in-noise tests. In this example, adaptive speech-in-noise testing reveals that the SRT for the hearing-impaired listener equals 0 dB (which is 5 dB worse than the SRT for normal-hearing listeners, depicted by the horizontal arrow). When using a fixed-SNR speech-in-noise test, the percentage-correct would be approximately 25% for an SNR of −3 dB (which is 50% less than the score for normal-hearing listeners, depicted by the vertical arrow). The arrows are also plotted in Fig. 1(A). The aim of this paper is to provide equations to convert between SRT and percentage-correct, given the speech recognition function of normal-hearing listeners. Thus, for the example above, a formula will be provided to convert the 0 dB SRT to 25% correct at SNR = –3 dB, but also for the percentages correct at any SNR, and vice versa.

Because the normal-hearing and hearing-impaired SII functions are linear and reach a value of zero at the same point (–15 dB SNR), each arrow in Fig. 1(A) fully defines the hearing-impaired SII function, and relatively simple relationships exist between the different measures (i.e., SRT and percentage-correct).

The speech recognition function, P(SNR), is a sigmoid function that can be approximated by, for example, the cumulative normal distribution or a logistic function. Note that, traditionally, the transfer function is described by $P(SII) = (1\text{–}10^{-SII/Q})^N$ (Fletcher and Galt, 1950; Studebaker and Sherbecoe, 1991), which yields a similar function for the speech recognition function with SII replaced by (SNR+15)/30 when using the previously mentioned approximation of the SII model. As indicated by Studebaker and Sherbecoe (1991), a number of equations could describe the transfer function and whatever gives the best fit may be best to use.

The logistic function is often used to describe the speech recognition function because of its ease of computation (e.g., Gordon-Salant and Fitzgibbons, 1995; Brand and Kollmeier, 2002; MacPherson and Akeroyd, 2014). Then the speech recognition function for normal-hearing listeners, called the reference speech recognition function, is described as

$$P(SNR) = \frac{100}{1 + \exp\left[-0.04 \cdot S_{50,NH}(SNR - SRT_{NH})\right]}, \quad (1)$$

where $S_{50,NH}$ is the slope of the normal-hearing speech recognition function at 50% correct, expressed in %/dB, and SRT_{NH} is the SRT for normal-hearing listeners. Slope values differ between stimuli sets (MacPherson and Akeroyd, 2014) and range between approximately 6%/dB for phonemes in words (Smits *et al.*, 2013) to 13% for high-predictability sentences (Versfeld *et al.*, 2000), up to 19% for digit triplets (Smits *et al.*, 2016). The different slopes of the SII functions for normal-hearing listeners and hearing-impaired listeners yield differences in slopes of the speech recognition functions. Note that the amount of speech information available to the listeners (represented by the SII value) is the same for all listeners at their SRT. The ratio of the slopes of the SII functions simply equals the ratio of the slopes of the speech recognition functions (Rhebergen and Versfeld, 2005; Smits and Festen, 2013). This results in the following equation for the slope of the speech recognition function for the hearing-impaired listener [Eq. (4) in Smits and Festen, 2011]:

$$S_{50,HI} = S_{50,NH}\frac{SRT_{NH} + 15}{SRT_{HI} + 15}. \quad (2)$$

When replacing $S_{50,NH}$ in Eq. (1) by the slope of the speech recognition function for the hearing-impaired listener [Eq. (2)], and replacing SRT_{NH} by SRT_{HI}, the speech recognition function for the hearing-impaired listener is obtained. The resulting equation describes the percentage-correct score for a hearing-impaired listener, P_{HI}, at a fixed SNR, SNR_{fixed}, given their SRT,

$$P_{HI} = \frac{100}{1 + \exp\left[-0.04 \cdot S_{50,NH} \frac{(SRT_{NH}+15)}{(SRT_{HI}+15)} (SNR_{fixed} - SRT_{HI})\right]}. \quad (3)$$

From Eq. (3), it is noticeable that parameter SRT_{HI} affects both slope and position of the speech recognition function, as shown in Fig. 1(B).

For the converse calculations where the percentage-correct score is known, the equivalent SRT_{HI} can be calculated using the following formula:

$$SRT_{HI} = \frac{(SRT_{NH} + 15) \cdot (SNR_{fixed} + 15)}{(SRT_{NH} + 15) - \frac{1}{0.04 \cdot S_{50,NH}} \ln\left[\frac{100}{P_{HI}} - 1\right]} - 15. \quad (4)$$

Equations (3) and (4) are valid for the conversion of speech recognition scores when a logistic function can approximate the normal-hearing speech recognition function. A general approach valid for any monotonic rising normal-hearing speech recognition function is given in the Appendix.

Figure 2 illustrates examples of conversion functions [Eq. (3) by choosing $SRT_{NH} = -5$ dB SNR and $S_{50,NH} = 14\%$/dB] for various values of SNR_{fixed}.

Figure 2 illustrates the nonlinear relationship between SRT and percentage-correct for all values of SNR_{fixed}. The floor and ceiling effects associated with fixed-SNR methods are reflected in the asymptotic behavior of the curves. The insensitivity of fixed-SNR methods to detect differences between listeners for certain combinations of SNR_{fixed} and SRT is illustrated in Fig. 2. For example, when testing listeners with SRTs in the range between −5 and 0 dB SNR at a fixed SNR of 5 or 10 dB, then the percentage-correct scores for these listeners are between 90% and 100%. Thus, a large number of stimuli are needed to detect significant differences between them. It may be concluded that for each fixed SNR value, only the results for listeners with SRTs within a range of approximately 5 dB provide valuable information (i.e., SRTs corresponding to the steep sloping part of the curves).

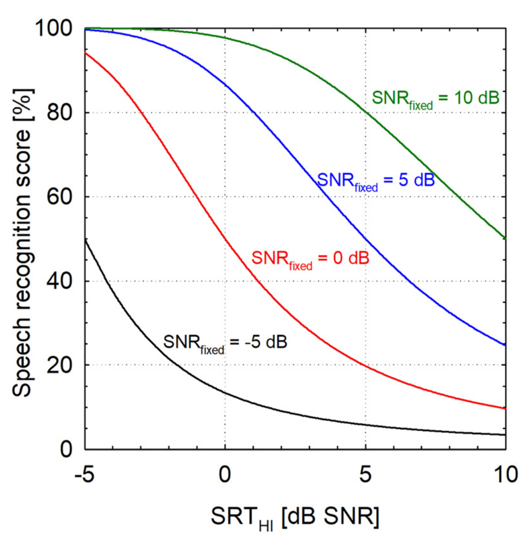

FIG. 2. (Color online) Speech recognition scores in percentage-correct for fixed-SNR methods as a function of the SRT_{HI} of the listener. Results for different values of SNR_{fixed} are shown.

III. VALIDATION OF THE CONVERSION METHOD: THE SII MODEL AS A BENCHMARK

To validate the accuracy of the conversion using the equations from the present study, we compared results from using the proposed simple analytical equations to results obtained from critical-frequency-band SII model calculations, which is currently the standard for calculating speech intelligibility. As indicated in the introduction, when the speech signal, noise signal, and hearing thresholds are known, then the SII model can calculate the amount of speech information available to the listener. These SII model calculations are used to determine the SRT and percentage-correct scores at various fixed SNRs for a representative set of "standard audiograms" from hearing-impaired listeners. Then, these percentage-correct scores are compared to percentage-correct scores calculated directly from the SRTs using Eq. (3).

The SII model is widely used to calculate the intelligibility of speech in steady-state noise. It provides a way of quantifying the effect of audibility on speech intelligibility and uses the listener's hearing threshold to account for the loss of audibility that hearing-impaired listeners may have. Several suggestions to incorporate the effects of suprathreshold deficits on speech intelligibility have been proposed to improve the model predictions for hearing-impaired listeners. The equations from the current paper are based on the fundamentals of the SII model. Essentially, the SII model was simplified by omitting self-speech masking, upward spread of masking, the level distortion factor, and by considering the loss of audibility due to an elevated hearing threshold constant. The latter means that it is assumed that the proportion of speech that is below the hearing threshold is constant within the range of presentation levels used in the tests (see Discussion section). The most notable difference between the two approaches is that the hearing thresholds and spectra of the speech and noise are needed for SII model calculations, whereas the equations in the current paper can be used without knowing the pure tone thresholds and the spectra of the stimuli.

A. Methods

1. Audiograms

A representative set of 60 typical audiograms from Bisgaard *et al.* (2010) was used. They constructed these audiograms from a large data set of recorded audiograms, and they cover the entire range of audiograms met in clinical practice. Bisgaard *et al.* (2010) present the hearing threshold for eight test frequencies for each of these 60 typical audiograms. The hearing thresholds for the 21 frequencies used in the SII critical-frequency-band procedure were based on these eight test frequencies.[2]

2. Reference speech recognition function

The reference (or normal-hearing) speech recognition function was based on the speech recognition function of the American English HINT (Nilsson *et al.*, 1994) and is represented by a logistic function [Eq. (1)] with $SRT_{NH} = -2.5$ dB (Eisenberg *et al.*, 1998) and $S_{50,NH} = 13.6$%/dB. This slope value represents an average steepness of 11.8%/dB between 20%-correct and 80%-correct as reported by Eisenberg *et al.* (1998).

3. SII model calculations

The critical-frequency-band method of the SII was used for the calculations. This method uses 21 frequency bands and is the most accurate method in the standard. The standard speech spectrum and the band importance function for SPIN were used in the calculations. It was assumed that the noise level was fixed at 65 dB SPL while the speech level varies, as is typically done when administering the HINT. The following steps were taken:

(a) The SII values for HINT sentences presented at SNRs ranging from −15 to 15 dB were calculated (0.1 dB steps), assuming hearing thresholds at 0 dB hearing level (HL). It also included the SII value corresponding to SRT_{NH}. Then, the transfer function was determined from the range of SII values and the reference speech recognition function.
Next, for each of the 60 standard audiograms:
(b) the SRT_{HI} that corresponds to the SII value for normal-hearing listeners was calculated
(c) SII values corresponding to $SNR_{fixed} = -5, 0, 5, 10$ dB were calculated, and
(d) percentage-correct for these fixed SNRs were calculated from (c) and the transfer function determined in (a).

These calculations produced for each audiogram (1) the SRT_{HI} and (2) percentage-correct scores for the series of fixed SNRs.

The calculations were also performed with a modified, more elaborated SII model, which incorporates the supra-threshold deficits that hearing-impaired listeners may have. In this modified model, an empirically derived hearing loss desensitization factor, proposed by Ching *et al.* (2015) and used in the NAL-NL2 fitting rule, was added to the SII critical-frequency-band method. The amount of desensitization for each frequency band is non-linearly dependent on the hearing threshold for that frequency band in this modified SII model.

4. Direct conversion of SRT to percentage correct with Eq. (3)

Equation (3) was used to convert the SRT_{HI}s directly to percentage-correct scores. Note that no hearing thresholds were used in this calculation but only the hearing-impaired SRT and the reference speech recognition function.

B. Results

Figure 3(A) shows the percentage-correct scores calculated directly from SRT_{HI} [Eq. (3)], thus without using the audiogram, against the percentage-correct scores determined from the SII model and the audiogram, for four fixed SNR levels. Figure 3(B) shows the results of these calculations for the modified SII model. Only data points with percentage-correct scores between 5% and 95% are shown because the estimates of percentage-correct scores are highly sensitive to errors in the SRT and the reference speech recognition function outside this range. Overall, there were very strong correlations between the scores ($r = 0.993$ and $r = 0.994$, respectively). The root-mean-square (RMS) errors in percentages for comparing the standard critical-frequency-band SII model, Fig. 3(A), and the modified SII model, Fig. 3(B), are approximately 4%. Separate calculations for the fixed-SNR groups show that RMS errors are largest for the highest SNR_{fixed} (i.e., 10 dB) and smallest for the lowest SNR_{fixed} (i.e., –5 dB).

IV. EXPERIMENTAL VALIDATION OF THE CONVERSION METHOD: AFRIKAANS SENTENCE-IN-NOISE TEST

The criterion validity of the conversion method test was assessed by comparing SRTs calculated from percentage-correct scores to SRTs measured with an adaptive procedure, which we consider the standard for determining the SRT. First, the reference speech recognition function was determined from an experiment with normal-hearing participants. Second, SRTs and percentage-correct scores at fixed SNRs were determined in a group of participants with different degrees of hearing loss. The SRTs were measured with a standard adaptive procedure and compared to SRTs which were determined by converting percentage-correct scores using our proposed method.

A. Methods

1. Participants

Two groups of listeners participated in the experiment: normal-hearing listeners to determine the reference speech recognition function and a group of participants with diverse hearing losses.

Eighteen normal-hearing listeners participated (2 male, 16 female). Mean age was 22 years (ranged from 19 to 26 years of age). The better ear was used in the speech-in-noise experiment. Pure-tone thresholds for these ears were equal to or better than 15 dB HL for all octave frequencies from 250 to 8000 Hz.

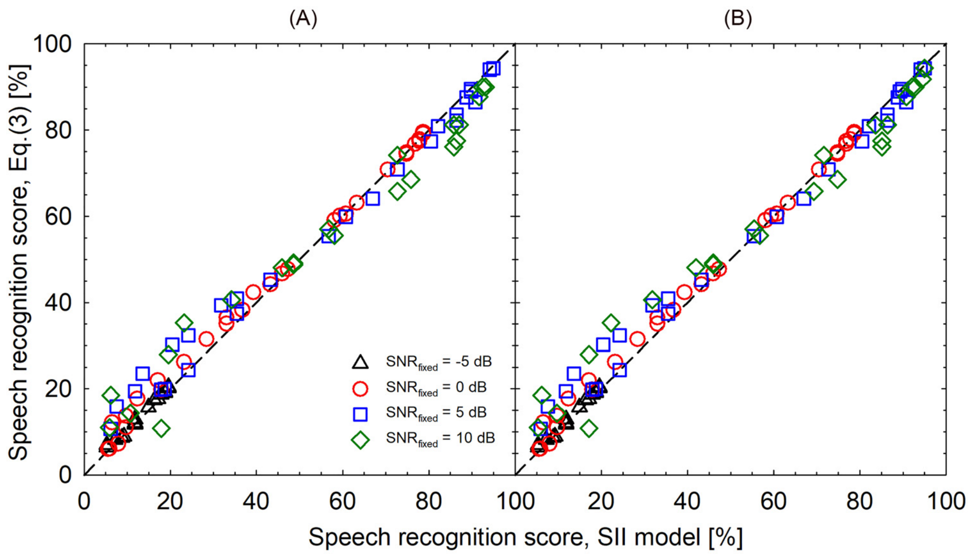

FIG. 3. (Color online) (A) Percentage-correct for four values of SNR_{fixed} calculated directly from the SRT_{HI} against percentage-correct determined from the SII model with the standard critical-frequency-band method and (B) similar calculations but for the modified SII model with a hearing loss desensitization factor.

The mean age of the second group of 28 participants (9 male, 19 female) was 62 years (ranged from 23 to 86 years of age). The pure tone thresholds of the participants' ears used in the experiment are shown in Fig. 4. The mean hearing loss, averaged across 500, 1000, 2000, and 4000 Hz, was 33 dB (standard deviation, SD = 15, ranged from 3 to 61).

The study was approved by the Faculty of Humanities Research Ethics Committee, University of Pretoria (Approval Number: GW20180717HS).

2. Stimuli

The target sentences were from the Afrikaans test for sentence recognition in noise (Theunissen *et al.*, 2011).

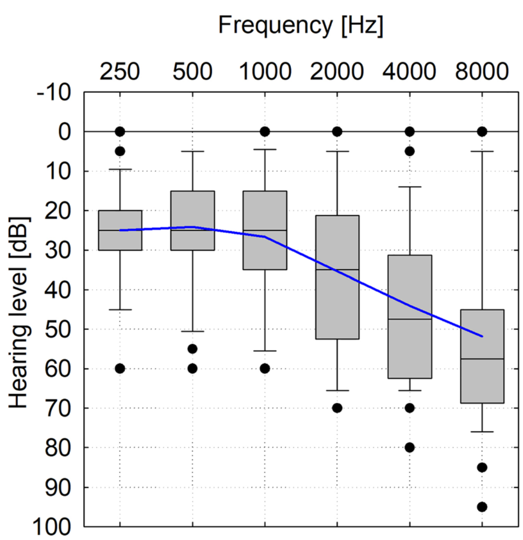

FIG. 4. (Color online) Audiograms (test ears) the group of for participants with different degrees of hearing loss. The boundaries of each box indicate the 25th and 75th percentiles, a line within the box marks the median. Whiskers indicate the 90th and 10th percentiles. The mean thresholds are denoted with a thick blue line.

These high-predictability sentences (e.g., "Die vrou het haar huis opgeruim" translated as "The woman cleaned her house") were spoken by a female talker and are similar to the HINT sentences for the English language (Nilsson *et al.*, 1994). Speech-shaped noise was constructed from Gaussian noise that was filtered to match the long-term power spectrum of the sentences. The test consists of 18 lists of ten sentences.

3. Procedure

The group of normal-hearing listeners was tested to determine the reference speech recognition function. The sentence lists were presented at eight different SNRs ranging from –10 to 4 dB in 2-dB steps. Among the 18 lists, eight were selected (one for each SNR) and presented to each participant. The sentence lists and order of the presentation levels were counterbalanced across the participants. Thus, each of the 180 sentences was presented once at each of the SNRs.

For the second group of participants, adaptive SRT tests were repeated twice for each participant, and four fixed-SNR tests were administered. No sentence list was presented more than once to any participant. A one-up one-down adaptive procedure with a fixed step size of 2 dB was used to determine the SRT (Theunissen *et al.*, 2011). The SRT was calculated by averaging the last 7 SNRs. The average of the two SRT estimates was used to choose the fixed-SNR levels. The two closest SNRs from the following SNRs were chosen: –5, 0, 5, and 10 dB SNR. For example, if one participant had an average SRT of 1.3 dB SNR, then the fixed SNRs of 0 and 5 dB were chosen. Two lists of ten sentences were presented at each of the two chosen SNRs.

For both the normal-hearing listeners and the second group of participants, the stimuli were presented monaurally at a fixed overall level of 65 dB SPL over Sennheiser

https://doi.org/10.1121/10.0005877

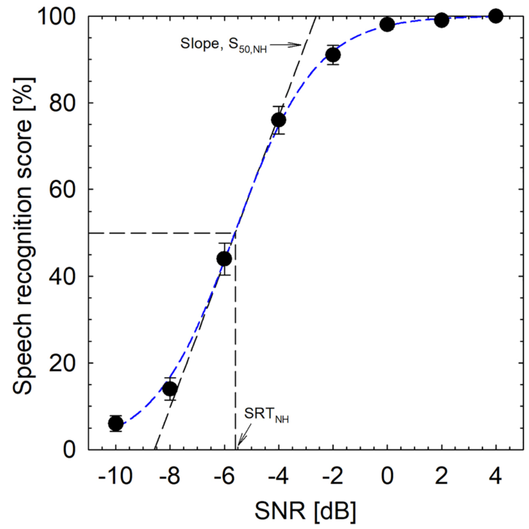

FIG. 5. (Color online) Reference speech recognition function for the Afrikaans sentence-in-noise test. The dashed line is a fitted logistic function with $S_{50,NH} = 16.8\%/\text{dB}$, and $SRT_{NH} = -5.6$ dB SNR.

HDA280 headphones in a sound-treated room. One adaptive practice list was provided before testing.

B. Results

1. Reference speech recognition function (normal-hearing listeners)

Figure 5 shows the mean percentage correct, averaged across participants and sentences, as a function of SNR. Each data point represents the mean and standard error of the mean of 180 presentations. The reference speech recognition function is depicted by the dashed line and was determined by fitting a logistic function [Eq. (1)] to the raw data. The fitting procedure provided estimations for the two parameters of the logistic function: $S_{50,NH} = 16.8\%/\text{dB}$, and $SRT_{NH} = -5.6$ dB SNR.

2. Converting percentage-correct to SRT (participants with different degrees of hearing loss)

For each participant from the second group, the two measured SRTs from the adaptive procedure were averaged to obtain $SRT_{adaptive}$, and the number of correctly recognized sentences from the 20 presentations at each fixed SNR was used to determine the percentage-correct score. Figure 6(A) shows the percentage-correct scores against $SRT_{adaptive}$. The predicted relationships between these parameters are depicted by dashed lines representing Eq. (3) with $S_{50,NH} = 16.8\%/\text{dB}$ and $SRT_{NH} = -5.6$ dB SNR.

The percentage correct scores were converted to $SRT_{fixed\,SNR}$ by using Eq. (4) with $S_{50,NH} = 16.8\%/\text{dB}$, and $SRT_{NH} = -5.6$ dB SNR. Five of the $SRT_{fixed\ SNR}$s could not be calculated because 0 or 100% of the sentences were correctly recognized. For a participant with severe hearing loss, $SRT_{adaptive}$ equaled 14.6 dB SNR, and the fixed-SNR procedure was applied only at SNR = 10 dB. Thus, a total of 50 $SRT_{fixed\ SNR}$ remained for analyses. There was no significant difference between $SRT_{adaptive}$ (M = 0.94, SD = 5.29) and $SRT_{fixed\ SNR}$ (M = 0.94, SD = 4.91), $t(49) = 0.015$, $p = 0.988$. The relationship between both SRT estimates is highlighted in Fig. 6(B). It shows $SRT_{fixed\ SNR}$ against $SRT_{adaptive}$. The dashed diagonal line represents equal performance. There is a strong correlation between $SRT_{adaptive}$ and $SRT_{fixed\ SNR}$, $r(48) = 0.866$, $p < 0.001$.

V. DISCUSSION

The importance of assessing speech recognition abilities in noise for hearing-impaired listeners is well recognized. Fixed-SNR methods and adaptive procedures both have their advantages and disadvantages. This paper provides formulae to convert fixed-SNR speech recognition scores, expressed in percentage-correct, to SRTs and vice versa, making a comparison of results possible. Importantly, neither the pure tone thresholds nor the speech and noise

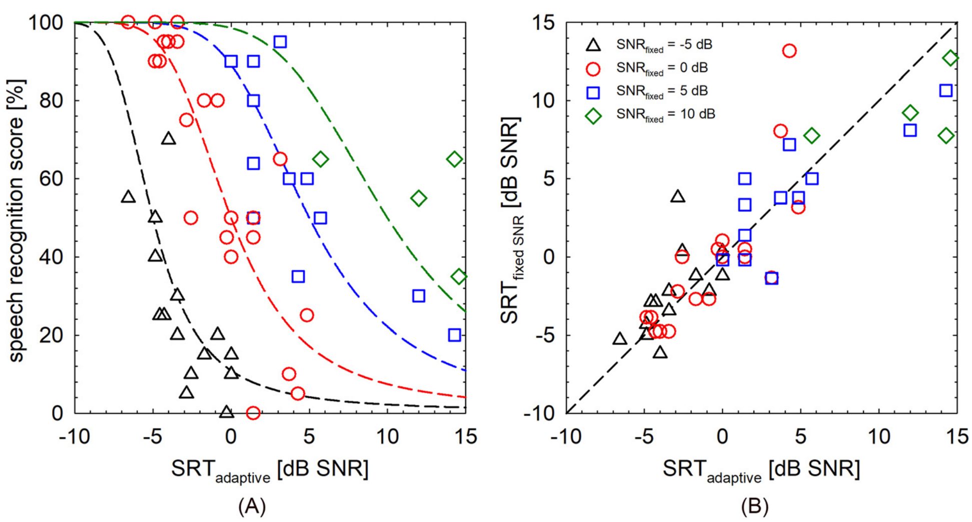

FIG. 6. (Color online) (A) Percentage correct scores at four different fixed SNR values against SRT. Dashed lines represent Eq. (3) with the two parameters from the reference speech recognition function. (B) SRTs calculated from percentage-correct scores at fixed SNRs against SRTs determined with a standard adaptive procedure.

spectra are needed for the conversion. The derived equations can be used for all speech materials when the normal-hearing speech recognition function for that specific speech material is known.

Reasons why the proposed method does not require thresholds and spectra information, in opposition to the SII model which does, include first the effect of elevated thresholds being similar for fixed and adaptive SNR procedures. In both procedures, they can cause inaudibility of portions of the speech, which results in a higher SRT or a lower percentage-correct score. Second, because it is required for our approach that the spectra from the speech and noise are equal, the relationship between SII values and SNR (the SII function) can be approximated very closely by a linear function.

Both testing methods have been in use for decades, and surprisingly, to our knowledge, no conversion methods between % and dB have been available to date. However, they aim to measure the same construct (i.e., the ability to recognize speech in noise) using the same stimuli and response task. If it is ensured that the fixed-SNR method and the adaptive procedure measure exactly the same construct, it should be possible to convert from one measure to another. This, however, is not always the case, which undermines the validity of the presented equations. The audibility component of the hearing loss could be different for both methods, which occurs when the speech levels are not equal, and more speech is audible in one of the methods. For example, when percentage-correct scores are measured at a low SNR, outside the range of SNRs used in the adaptive procedure, a larger proportion of the speech will be inaudible. This means that the ability to recognize speech in noise is different for both measurements conditions, and both tests measure a slightly different construct. Consequently, the conversion formulae will not be fully correct. This effect is clearly demonstrated for the $SNR_{fixed} = 10$ dB condition in Fig. 3(A). The direct calculations underestimate percentage-correct scores for high percentages and overestimate percentage-correct scores for low percentages. Listeners with percentage-correct scores between 50% and 90% at the relatively favourable SNR_{fixed} of 10 dB have relatively poor audiograms (i.e., high pure-tone thresholds). Their SRTs are lower than 10 dB because their percentage-correct scores are higher than 50%. When measuring at a fixed SNR of 10 dB, a larger proportion of the speech signal will be audible during the adaptive procedure where lower SNRs are used. The SII model takes the improved audibility into account, but the direct calculations do not. For listeners with percentage-correct scores between 10% and 50%, the audibility reduces when measuring at a fixed SNR of 10 dB because these listeners have SRTs higher than 10 dB.

Figures 2 and 6(A) illustrate the caution required when interpreting speech recognition data in noise. Percentage-correct scores from fixed-SNR methods are particularly vulnerable to misinterpretation because of floor and ceiling effects and because percent-correct scores cannot be compared when obtained at different SNRs.

The proposed method for converting percentage-correct scores to SRTs was experimentally validated (Sec. IV) for a standard speech-in-noise test with Afrikaans high predictability sentences. The reference speech recognition function could be described accurately with a two-parameter logistic function, and the subsequent experiment revealed that SRTs could be predicted very well from percent-correct scores using the two parameters from the reference speech recognition function [Fig. 6(B)]. The scatter around the line of equality can be attributed, at least for the most part, to the measurement error because the correlation between both values for $SRT_{adaptive}$ [$r(26) = 0.882$, $p < 0.001$] is similar to the correlation between $SRT_{adaptive}$ and the SRT which was derived from the percentage-correct score.

One of our methods important outcomes is that it can quantify differences in speech recognition even from previous experiments or published data if the reference speech recognition function is known. For example, from our experimental data (Sec. IV), the average number of correctly recognized sentences equals 27.9% at -5 dB (14 participants), and the average number of correctly recognized sentences at 0 dB equals 62.6% (13 participants). The difference in the ability to recognize sentences in noise between these groups can be quantified by converting these mean percentage-correct scores to mean SRTs and calculating the difference in dB. It yields a mean difference of 6.7 dB, which is very similar to the value of 7.3 dB for the mean difference between the SRTs from the adaptive procedure. Although the number of data points is low for the highest SNR (i.e., 10 dB SNR), the data in Fig. 6 indicates that for this SNR, the SRTs derived from percentage-correct scores are lower than the SRTs from the adaptive procedure. A likely reason for this is the bias in SRT estimates for participants with severe hearing losses who do not reach 100% correct scores in quiet. Then the adaptive procedure does not work effectively, and the target point differs from 50% (Smits and Houtgast, 2006).

An important aspect of a speech-in-noise test is the precision, represented by the measurement error. We used a brute-force method (Smits and Houtgast, 2006) to determine the measurement error for the adaptive procedure from the Afrikaans sentence-in-noise test. For the standard ten-sentences adaptive procedure, the calculated measurement error ranges from 1.2 dB for normal-hearing listeners to 2.3 dB for listeners with an SRT of 10 dB. The measurement error reduces with a factor $1/\sqrt{2}$ to, respectively, 0.8 and 1.6 dB when the average of two SRT estimates is used as in Fig. 6(B). The measurement error for the fixed-SNR procedure depends on the probability of a correct response, p, and the number of presentations (Thornton and Raffin, 1978). Because the slope of the speech recognition function is very shallow for high recognition probabilities, the measurement error for SRTs that are calculated from percentage-correct scores becomes large as well when based on these probabilities. However, even when percent-correct scores are not too low or too high (approximately between 10% and 90%), the measurement error of the SRTs derived from these percent-

https://doi.org/10.1121/10.0005877

correct scores is larger than the measurement errors from the adaptive procedure.

Besides the random error caused by the limited precision of the measurement, converting one measure to another will introduce a systematic error because the reference speech recognition function is not known exactly. We explored the effect of inaccurate values of $S_{50,NH}$ and SRT_{NH} on percentage-correct scores calculated from SRTs [Eq. (3)] by using the partial derivative of P with respect to $S_{50,NH}$ and SRT_{NH}. The standard errors of $S_{50,NH}$ and SRT_{NH} for the reference speech recognition function for the Afrikaans test for sentence recognition in noise (Sec. IV B 1) are 0.9%/dB and 0.6 dB SNR, respectively. The effect of the calculated percentage-correct scores is approximately between −1 and 1% for both errors when varying of SRTs between −5 to 15 dB and SNR_{fixed} between −5 and 10 dB. The effect is positive when $SRT_{HI} < SNR_{fixed}$, negative when $SRT_{HI} > SNR_{fixed}$, and 0% when $SRT_{HI} = SNR_{fixed}$. The effect on the calculated percentage-correct score is relatively small compared to the effect of the error in a standard SRT estimate, but this emphasizes the importance of accurately determining the reference speech recognition function.

The formulae in this paper for converting the results from speech-in-noise measurements apply only when the speech-in-noise test meets several requirements: (1) the steady-state masking noise has exactly the same spectrum as the long term speech spectrum; (2) the SNR is correctly defined;[3] (3) the steady-state noise is Gaussian noise; and (4) bandwidth, and quality of the signals are identical for all measurements.[4] The derivations in this paper assume that the SII function for hearing-impaired listeners starts at –15 dB SNR, as for normal-hearing listeners. It seems reasonable to expect a higher value than –15 dB SNR for some hearing-impaired listeners, particularly cochlear implant users (Smits and Festen, 2011). However, the approach used here follows the basic assumptions of the SII model and the modified SII model (Ching *et al.*, 2015) and these models both assume that the speech information becomes accessible for SNRs above –15 dB for normal-hearing and hearing-impaired listeners. This assumption proved to be true for group data (Smits and Festen, 2011), but may overlook individual differences between hearing-impaired listeners.

It is possible to derive formulae that could be used for tests that do not meet all these requirements, but these formulae will be more complex. The SII function will be shallower and will not start from –15 dB SNR when, for instance, white noise is used for masking instead of noise that matches the spectrum of the speech material. Many speech-in-noise tests use fluctuating noises like modulated speech noise, single-talker interference, or multitalker babble. The SII function for a fluctuating noise is shallower and stretches over a larger range than the 30 dB range for steady-state noise (Smits and Festen, 2013). Formulae to convert percentage-correct to SRT or vice versa can be derived when the exact SII function for the fluctuating noise is known. Note that the terms steady-state noise and fluctuating noise are used here in the traditional sense, where steady-state noise refers to Gaussian noise shaped to match the spectrum of the speech. In earlier work, it has been assumed that steady-state noise produces energetic masking, which limits speech recognition. However, Stone *et al.* (2012) and Stone and Moore (2014) demonstrated that the inherent fluctuations in the noise primarily produce the masking in this type of noise. They use the term notionally steady-state noise to take into account this finding.

VI. CONCLUSIONS

We have derived a simple analytical method to convert between results from fixed-SNR procedures, expressed in percentage-correct, and SRTs from adaptive procedures, expressed in dB SNR, when the speech-in-noise test uses steady-state masking noise that matches the spectrum of the speech. Pure tone thresholds are not required for these calculations, which makes the method widely applicable. The experimental data and comparison to the SII model demonstrate the validity of the approach and the usefulness of the proposed equations. These equations permit an easy comparison of results from different studies on speech recognition in noise, to quantify differences in speech recognition abilities in noise when measurements are performed at different SNRs, and to compare percentage-correct scores to SRTs.

APPENDIX: DERIVATION OF GENERAL EQUATIONS

Figure 7(A) shows the SII function for normal-hearing listeners and an example SII function for a hearing-impaired listener. The corresponding speech recognition functions are shown in Fig. 7(B). Assume that the SRT of the hearing-impaired listener, $SRT_{HI,}$ has been measured but the percentage-correct for this hearing-impaired listener at a certain fixed SNR, SNR_{fixed}, needs to be known. The lower open circle represents the unknown percentage correct for the hearing-impaired listener at SNR_{fixed}. The SNR where a normal-hearing listener reaches the same percentage-correct (denoted by $SNR_{NH,P_{HI}}$) is also shown in Fig. 7(B).

These points are also shown in Fig. 7(A) where they fall on the linear SII functions. Then geometry shows

$$\frac{(SNR_{NH,P_{HI}} + 15)}{(SNR_{fixed} + 15)} = \frac{(SRT_{NH} + 15)}{(SRT_{HI} + 15)}. \quad \text{(A1)}$$

Hence,

$$SNR_{NH,P_{HI}} = (SNR_{fixed} + 15) \cdot \frac{(SRT_{NH} + 15)}{(SRT_{HI} + 15)} - 15, \quad \text{(A2)}$$

which is the SNR where normal-hearing listeners reach the same percentage-correct as a hearing-impaired listener at SNR_{fixed}. The corresponding percentage-correct score can be derived from the normal-hearing speech recognition function, $P_{NH}(SNR)$, with the SNR from Eq. (A2), which yields

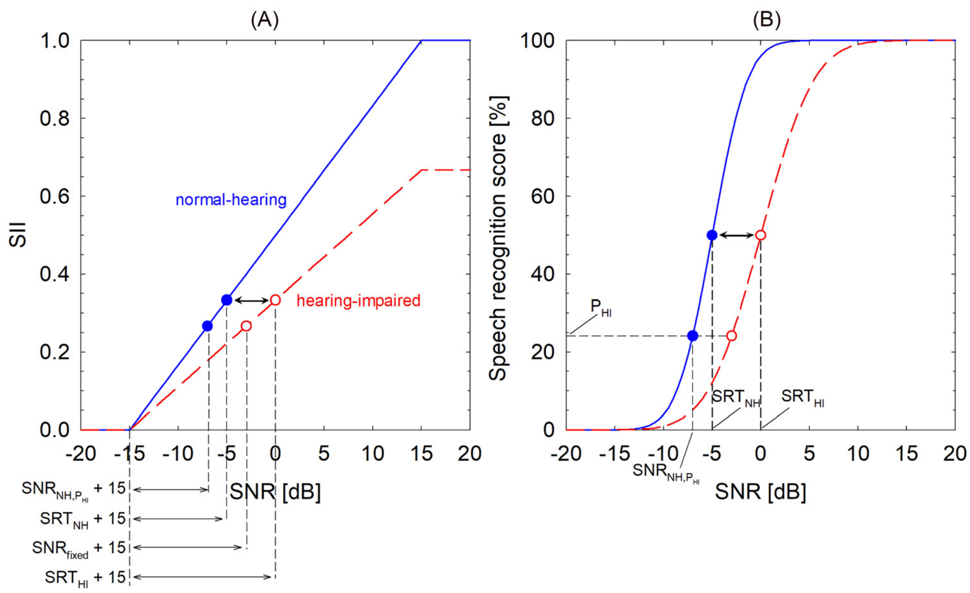

FIG. 7. (Color online) (A) SII functions for normal-hearing listeners and a hearing-impaired listener. (B) Speech recognition functions for these listeners. The black arrows show the difference in SRT between normal-hearing listeners and a hearing-impaired listener. See the main text for further explanation.

$$\mathrm{P}_{\mathrm{HI}} = \mathrm{P}_{\mathrm{NH}}\left((\mathrm{SNR}_{\mathrm{fixed}} + 15)\cdot\frac{\mathrm{SRT}_{\mathrm{NH}} + 15}{\mathrm{SRT}_{\mathrm{HI}} + 15} - 15\right), \quad \text{(A3)}$$

The normal-hearing speech recognition function $\mathrm{P}_{\mathrm{NH}}(\mathrm{SNR})$ is a sigmoid function and determined by fitting a function to data points that are usually measured at fixed SNRs (see Sec. IV). As mentioned in the main text, both cumulative normal distributions, logistic functions, and the function proposed by Fletcher and Galt (1950) may be used. These functions have two free fitting parameters, and, as the example in Fig. 8 shows, the correlation between measured and fitted values is excellent for all functions.

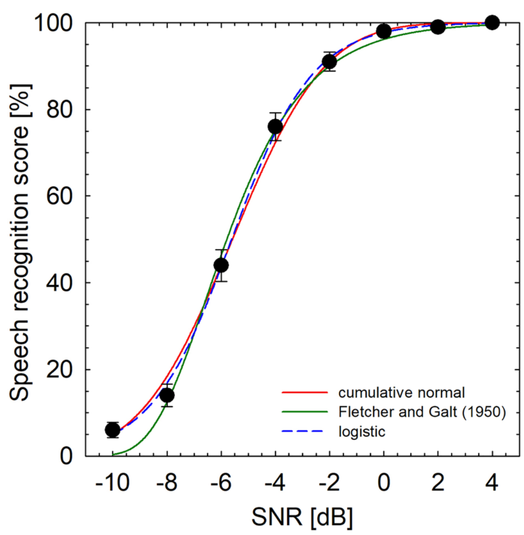

FIG. 8. (Color online) Speech recognition data from normal-hearing listeners (from Sec. IV). A logistic function, cumulative normal distribution, and the function proposed by Fletcher and Galt (1950) were fitted to the data. All functions demonstrate excellent agreement with the data ($R^2 = 0.999$, $R^2 = 0.997$ and $R^2 = 0.997$, respectively).

For the converse calculations where the percentage-correct score at a certain fixed SNR is known, and the equivalent SRT needs to be calculated, Eq. (A3) can be rewritten. First, the inverse normal-hearing speech recognition function, $\mathrm{P}_{\mathrm{NH}}^{-1}$, is used, which yields

$$\mathrm{P}_{\mathrm{NH}}^{-1}(\mathrm{P}_{\mathrm{HI}}) = (\mathrm{SNR}_{\mathrm{fixed}} + 15)\cdot\frac{\mathrm{SRT}_{\mathrm{NH}} + 15}{\mathrm{SRT}_{\mathrm{HI}} + 15} - 15. \quad \text{(A4)}$$

This can be rewritten as

$$\mathrm{SRT}_{\mathrm{HI}} = \frac{(\mathrm{SRT}_{\mathrm{NH}} + 15)\cdot(\mathrm{SNR}_{\mathrm{fixed}} + 15)}{\mathrm{P}_{\mathrm{NH}}^{-1}(\mathrm{P}_{\mathrm{HI}}) + 15} - 15. \quad \text{(A5)}$$

Thus, the $\mathrm{SRT}_{\mathrm{HI}}$ (i.e., SRT for a hearing-impaired listener) can be calculated given the percentage-correct, P_{HI}, for that hearing-impaired listener measured at a certain fixed SNR and the normal-hearing speech recognition function.

The formulae in the main text as well as in the formulae in the appendix are valid for converting percentage-correct to SRT and vice versa. However, sometimes a different target point is used and the SNR corresponding to x% correct, SNR_x, is determined. The formulae to convert SNR_x to SRT and vice versa are similar to Eqs. (A1) and (A2) and can be used before or after converting percentage-correct and SRT,

$$\mathrm{SRT}_{\mathrm{HI}} = (\mathrm{SRT}_{\mathrm{NH}} + 15)\cdot\frac{(\mathrm{SNR}_{x,\mathrm{HI}} + 15)}{(\mathrm{SNR}_{x,\mathrm{NH}} + 15)} - 15, \quad \text{(A6)}$$

$$\mathrm{SNR}_{x,\mathrm{HI}} = (\mathrm{SNR}_{x,\mathrm{NH}} + 15)\cdot\frac{(\mathrm{SRT}_{\mathrm{HI}} + 15)}{(\mathrm{SRT}_{\mathrm{NH}} + 15)} - 15. \quad \text{(A7)}$$

[1]Other procedures like the Spearman-Kärber procedure or maximum likelihood procedures can be used as well but they are less common in clinical speech-in-noise testing.

https://doi.org/10.1121/10.0005877

[2]See supplementary material at https://www.scitation.org/doi/suppl/10.1121/10.0005877 for the hearing thresholds in dB HL for the 60 typical audiograms.
[3]The definition of the SNR is not trivial for speech material, especially not for single words or digit-triplets because the duration of the silences between words influence the average speech level.
[4]This means for instance that identical headphones should be used for all measurements.

ANSI (**1986**). S3. 5-1969 (R1986), *American National Standard Methods for the Calculation of the Articulation Index* (ANSI, New York).

ANSI (**1997**). S3. 5-1997, *Methods for the Calculation of the Speech Intelligibility Index* (ANSI, New York), pp. 90–119.

Bench, J., Kowal, Å., and Bamford, J. (**1979**). "The BKB (Bamford-Kowal-Bench) sentence lists for partially-hearing children," Br. J. Audiol. **13**, 108–112.

Bilger, R. C., Nuetzel, J. M., Rabinowitz, W. M., and Rzeczkowski, C. (**1984**). "Standardization of a test of speech perception in noise," J. Speech Lang. Hear. Res. **27**, 32–48.

Bisgaard, N., Vlaming, M. S., and Dahlquist, M. (**2010**). "Standard audiograms for the IEC 60118-15 measurement procedure," Trends Amplif. **14**, 113–120.

Brand, T., and Kollmeier, B. (**2002**). "Efficient adaptive procedures for threshold and concurrent slope estimates for psychophysics and speech intelligibility tests," J. Acoust. Soc. Am. **111**, 2801–2810.

Brungart, D. S., Walden, B., Cord, M., Phatak, S., Theodoroff, S. M., Griest, S., and Grant, K. W. (**2017**). "Development and validation of the Speech Reception in Noise (SPRINT) Test," Hear. Res. **349**, 90–97.

Carhart, R., and Tillman, T. W. (**1970**). "Interaction of competing speech signals with hearing losses," Arch. Otolaryngol. **91**, 273–279.

Ching, T. Y., Dillon, H., and Byrne, D. (**1998**). "Speech recognition of hearing-impaired listeners: Predictions from audibility and the limited role of high-frequency amplification," J. Acoust. Soc. Am. **103**, 1128–1140.

Ching, T. Y., Quar, T. K., Johnson, E. E., Newall, P., and Sharma, M. (**2015**). "Comparing NAL-NL1 and DSL v5 in hearing aids fit to children with severe or profound hearing loss: Goodness of fit-to-targets, impacts on predicted loudness and speech intelligibility," J. Am. Acad. Audiol. **26**, 260–274.

Eisenberg, L. S., Dirks, D. D., Takayanagi, S., and Martinez, A. S. (**1998**). "Subjective judgments of clarity and intelligibility for filtered stimuli with equivalent speech intelligibility index predictions," J. Speech Lang. Hear. Res. **41**, 327–339.

Fletcher, H., and Galt, R. H. (**1950**). "The perception of speech and its relation to telephony," J. Acoust. Soc. Am. **22**, 89–151.

García-Pérez, M. A. (**1998**). "Forced-choice staircases with fixed step sizes: Asymptotic and small-sample properties," Vis. Res. **38**, 1861–1881.

Gilbert, J. L., Tamati, T. N., and Pisoni, D. B. (**2013**). "Development, reliability, and validity of PRESTO: A new high-variability sentence recognition test," J. Am. Acad. Audiol. **24**, 026–036.

Gordon-Salant, S., and Fitzgibbons, P. J. (**1995**). "Comparing recognition of distorted speech using an equivalent signal-to-noise ratio index," J. Speech Lang Hear. Res. **38**, 706–713.

Killion, M. C., Niquette, P. A., Gudmundsen, G. I., Revit, L. J., and Banerjee, S. (**2004**). "Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners," J. Acoust. Soc. Am. **116**, 2395–2405.

Kollmeier, B., Warzybok, A., Hochmuth, S., Zokoll, M. A., Uslar, V., Brand, T., and Wagener, K. C. (**2015**). "The multilingual matrix test: Principles, applications, and comparison across languages: A review," Int. J. Audiol. **54**, 3–16.

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

MacPherson, A., and Akeroyd, M. A. (**2014**). "Variations in the slope of the psychometric functions for speech intelligibility: A systematic survey," Trends Hear. **18**, 233121651453772.

Nilsson, M., Soli, S. D., and Sullivan, J. A. (**1994**). "Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise," J. Acoust. Soc. Am. **95**, 1085–1099.

Owen, J. H. (**1981**). "Influence of acoustical and linguistic factors on the SPIN test difference score," J. Acoust. Soc. Am. **70**, 678–682.

Pavlovic, C. V., Studebaker, G. A., and Sherbecoe, R. L. (**1986**). "An articulation index based procedure for predicting the speech recognition performance of hearing-impaired individuals," J. Acoust. Soc. Am. **80**, 50–57.

Plomp, R. (**1978**). "Auditory handicap of hearing impairment and the limited benefit of hearing aids," J. Acoust. Soc. Am. **63**, 533–549.

Rhebergen, K. S., and Versfeld, N. J. (**2005**). "A speech intelligibility index-based approach to predict the speech reception threshold for sentences in fluctuating noise for normal-hearing listeners," J. Acoust. Soc. Am. **117**, 2181–2192.

Ricketts, T., Bentler, R. A., and Mueller, H. G. (**2019**). *Essentials of Modern Hearing Aids: Selection, Fitting, and Verification* (Plural Publishing, San Diego, CA).

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

Smits, C., Watson, C. S., Kidd, G. R., Moore, D. R., and Goverts, S. T. (**2016**). "A comparison between the Dutch and American-English digits-in-noise (DIN) tests in normal-hearing listeners," Int. J. Audiol. **55**, 358–365.

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

Stone, M. A., Füllgrabe, C., and Moore, B. C. (**2012**). "Notionally steady background noise acts primarily as a modulation masker of speech," J. Acoust. Soc. Am. **132**, 317–326.

Stone, M. A., and Moore, B. C. (**2014**). "On the near non-existence of 'pure' energetic masking release for speech," J. Acoust. Soc. Am. **135**, 1967–1977.

Studebaker, G. A., and Sherbecoe, R. L. (**1991**). "Frequency-importance and transfer functions for recorded CID W-22 word lists," J. Speech Lang. Hear. Res. **34**, 427–438.

Theunissen, M., Hanekom, J. J., and Swanepoel, D. (**2011**). "The development of an Afrikaans test for sentence recognition thresholds in noise," Int. J. Audiol. **50**, 77–85.

Thornton, A. R., and Raffin, M. J. (**1978**). "Speech-discrimination scores modeled as a binomial variable," J. Speech Hear. Res. **21**, 507–518.

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

Wilson, R. H. (**2003**). "Development of a speech-in-multitalker-babble paradigm to assess word-recognition performance," J. Am. Acad. Audiol. **14**, 453–470.

Zekveld, A. A., Kramer, S. E., and Festen, J. M. (**2011**). "Cognitive load during speech perception in noise: The influence of age, hearing loss, and cognition on the pupil response," Ear Hear. **32**, 498–510.