

The Effect of Stimulus Audibility on the Relationship between Pure-Tone Average and Speech Recognition in Noise Ability

DOI: 10.3766/jaaa.19031

Andrew J. Vermiglio*
Sigfrid D. Soli†
Daniel J. Freed‡
Xiangming Fang§

Abstract

Background: The literature presents conflicting reports on the relationship between pure-tone threshold average and speech recognition in noise ability.

Purpose: The purpose of this retrospective study and meta-analysis was to determine the effect of stimulus audibility on the relationship between speech recognition in noise ability and bilateral pure-tone average (BPTA).

Research Design: Pure-tone threshold and Hearing in Noise Test (HINT) data from two data sets were evaluated. The HINT data from both data sets were divided into groups with complete and partial audibility of the HINT stimuli delivered at 65 dBA.

Study Sample: Normal and hearing-impaired participants were included in this retrospective study. For data set 1 ($n = 215$), a relatively weak relationship had been found between HINT thresholds and BPTA. For data set 2 ($n = 55$), a relatively strong relationship had been found between HINT thresholds and BPTA. For data set 1, only 10% of the participants had partial audibility of the HINT stimuli. For data set 2, 16% of the participants had partial audibility of the HINT stimuli.

Data Collection and Analysis: Pure-tone thresholds and HINT data were obtained from published and unpublished studies. HINT data were collected in a simulated soundfield environment under headphones using the standard HINT protocol. Statistical analyses included descriptive statistics, correlations, and a two-way analysis of variance (ANOVA), and multiple regression.

Results: A two-way ANOVA followed by post hoc analyses revealed a greater difference between the data sets for the Noise Front thresholds obtained with partial rather than complete audibility of the stimuli. A weak and nonsignificant relationship was found between BPTA_(0.5, 1.0, 2.0, 3.0, 6.0 kHz) versus HINT Noise Front thresholds for complete audibility data ($r = 0.060$, $p = 0.356$) and a strong relationship was found for the partial audibility data ($r = 0.863$, $p < 0.001$).

Conclusions: The proportion of partial audibility data in a given data set may influence the relative strength of the relationship between BPTA and HINT Noise Front thresholds. This brings into question the convention of using pure-tone average as a predictor of speech recognition in noise ability.

Key Words: gold-standard, pure-tone thresholds, speech recognition in noise ability

Abbreviations: AAO-ACO = American Academy of Otolaryngology, American Council of Otolaryngology; AMA = American Medical Association; ANOVA = analysis of variance; BKB-SIN = Bamford-Kowal-Bench Speech-in-Noise; BPTA = bilateral pure-tone average; HINT = Hearing in Noise Test; PTA = pure-tone average; QuickSIN = Quick Speech-in-Noise; SIN = Speech-in-Noise; SNR = signal-to-noise ratio; SPIN = Speech Perception in Noise; WHO = World Health Organization; WIN = Words-in-Noise

*Department of Communication Sciences and Disorders, East Carolina University, Greenville, NC; †House Clinic Los Angeles, CA; ‡Advanced Bionics LLC, Valencia, CA; §Department of Biostatistics, East Carolina University, Greenville, NC

Corresponding author: Andrew J. Vermiglio, Department of Communication Sciences and Disorders, East Carolina University, Greenville, NC 27834; Email: vermiglio@ecu.edu

INTRODUCTION

The Evaluation of Hearing Impairment

According to the Committee on the Conservation of Hearing of the American Academy of Ophthalmology and Otolaryngology (Lierle, 1959), hearing impairment should be evaluated in terms of the ability to hear “everyday speech under everyday conditions.” They also noted that “because of present limitations of speech audiometry, the hearing level for speech should be estimated from measurements with a pure-tone audiometer. For this purpose, the subcommittee recommends the simple average of the hearing levels at the three frequencies 500, 1000 and 2000 (Hz).” According to the Hearing Handicap Guide published by the American Academy of Otolaryngology and the American Council of Otolaryngology (AAO-ACO, 1979), a handicap or impairment is “a medical condition that affects one’s personal efficiency in the activities of daily living.” The AAO-ACO wrote that the basis for the calculation of hearing handicap should be modified to reflect the understanding of speech, not only in a quiet environment but also in the presence of “some noise.” At that time, they noted there was no standardized test for this assessment; therefore, they recommended the determination of hearing impairment based on the decibel sum of pure-tone threshold levels for 500, 1000, 2000, and 3000 Hz.

Dobie (2011) noted that the method for the determination of “hearing handicap” found in AAO-ACO (1979) is still recommended by the American Medical Association (AMA, 2008). Dobie evaluated the assumptions in AAO-ACO (1979) by determining the relationship between a “gold standard for self-assessed hearing disability” using the communication profile for the hearing-impaired (Demorest and Erdman, 1987) versus

pure-tone average (PTA) for 1,001 patients. Dobie reported modest correlations between $PTA_{(0.5, 1.0, 2.0, 3.0 \text{ kHz})}$ versus communication profile results for the better ear ($r = -0.385$) and the worse ear ($r = -0.282$). He concluded that the study supported the continued use of the 1979 AMA method for the determination of hearing impairment.

The Relationship between PTA and Speech Recognition in Noise Ability

Recall that the AAO-ACO (1979) stated that the estimation of hearing impairment should be based not only on the perception of speech in quiet but also in the presence of noise. A strong relationship between speech recognition in noise ability and PTA would validate the PTA as a reasonable measure of the perception of “everyday speech in everyday listening conditions.” Figure 1 shows the correlation coefficients for the relationships between $PTA_{(0.5, 1.0, 2.0 \text{ kHz})}$ or $PTA_{(0.5, 1.0, 2.0, 4.0 \text{ kHz})}$ versus speech recognition in noise performance across a number of studies (Lyregaard, 1982; Smoorenburg, 1992; Tschopp and Zust, 1994; Saunders et al, 2004; Smits et al, 2004; Vermiglio, 2007; Wilson, McArdle, et al, 2007; Vermiglio and Soli, 2008; Vermiglio et al, 2012). The strong relationships found in the studies by Saunders et al (2004) and Vermiglio and Soli (2008) imply that PTA may be used to predict speech recognition in noise ability. However, the weaker relationships found in the studies by Wilson, McArdle, et al (2007) and Vermiglio et al (2012) suggest that PTA is a poor predictor of speech perception in noise (SPIN) ability. The inconsistencies found across studies may be attributed to the differences in test protocols and/or participant characteristics.

For the studies shown in Figure 1, the numbers of participants ranged from 33 (Tschopp and Zust, 1994) to 215 (Vermiglio et al, 2012). Most studies included participants with normal pure-tone sensitivity and participants with hearing impairment. At least two studies included participants with conductive hearing losses (Lyregaard, 1982; Smits et al, 2004). Most studies included participants over the age of 55 years. The speech recognition in noise tests used in these studies included the Hearing in Noise Test (HINT; Nilsson et al, 1994; Vermiglio, 2008), the Bamford–Kowal–Bench Speech-in-Noise (BKB-SIN) test (Niquette et al, 2003), the Quick Speech-in-Noise (QuickSIN) test (Killion et al, 2004), the Words-in-Noise (WIN) test (Wilson, Carnell, et al, 2007), the German version of the SPIN test (Zust and Tschopp, 1993), the speech recognition in noise test developed by Plomp and Mimpen (1979), and a phoneme intelligibility in noise test (Lyregaard, 1982).

Most studies presented in Figure 1 used headphones for a monaural delivery of the stimuli. In four of the studies, the HINT was presented with (Saunders et al, 2004; Vermiglio and Soli, 2008; Vermiglio et al,

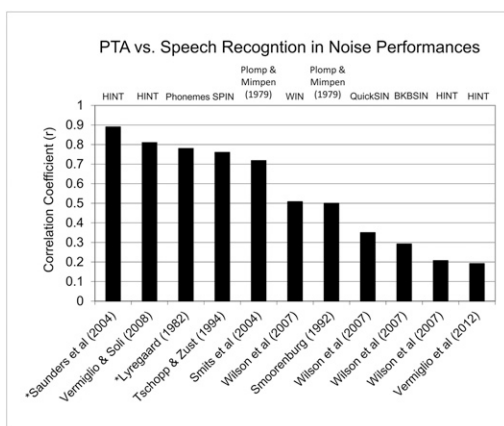


Figure 1. Correlation coefficients for the relationship between PTA and speech recognition in noise performance. $PTA_{(0.5, 1.0, 2.0 \text{ kHz})}$ was used for all of the studies shown with the exception of two. * = $PTA_{(0.5, 1.0, 2.0, 4.0 \text{ kHz})}$.

2012) or without (Wilson, McArdle, et al, 2007) a simulated soundfield environment under headphones. A few tests used words (modified rhyme and WIN tests). All other tests used sentences. Sentences are considered representative of “everyday speech” (Lierle, 1959; AAO-ACO, 1979). The masking noise used across studies was speech-shaped noise or multitalker babble. For most studies, the level of the speech was varied. The range of noise levels was from 65 dBA (e.g., Smoorenburg, 1992) to 100 dB SPL (Wilson, McArdle, et al, 2007). Most studies included monaural PTA. The studies by Vermiglio and colleagues used bilateral PTAs (BPTA) where the PTA for each ear was averaged together. The AMA initially used $PTA_{(0.5, 1.0, 2.0 \text{ kHz})}$ for the determination of hearing impairment (Lierle, 1959). Currently, the AMA uses $PTA_{(0.5, 1.0, 2.0, 3.0 \text{ kHz})}$ for this purpose (AMA, 2008). The World Health Organization (WHO) is currently using $PTA_{(0.5, 1.0, 2.0, 4.0 \text{ kHz})}$ for the determination of hearing impairment (Mathers et al, 2000). For the studies in Figure 1, all of the correlation coefficients are positive, indicating that a higher (poorer) PTA is associated with a more positive (poorer) signal-to-noise ratio (SNR) at threshold.

Although the variety in test parameters may explain some of the variability across outcomes, the studies with the strongest (Saunders et al, 2004; Vermiglio et al, 2012) and weakest (Wilson, McArdle, et al, 2007; Vermiglio and Soli, 2008) correlations all used the HINT. One explanation for some of the differences found across HINT studies is the differences in protocols. For example, Saunders et al (2004) used the simulated soundfield version of the HINT under headphones and averaged the thresholds for the noise right and noise left binaural listening conditions where the noise was presented at 90° and -90° , respectively, and the target speech was presented at 0° . The noise in this study was presented at 65 dBA. Wilson, McArdle, et al (2007), on the other hand, presented the HINT stimuli monaurally via head-

phones, and the noise was delivered at a nonstandard level of 90 dB SPL.

Only two of the studies, Vermiglio et al (2012) and Vermiglio and Soli (2008), used the same standard HINT protocol. However, the relationships found between speech recognition in noise thresholds and PTA were dissimilar across these two studies. The data from Vermiglio and Soli (2008) revealed a relatively strong relationship ($r = 0.810, p < 0.01$) and the data from Vermiglio et al (2012) revealed a relatively weak relationship ($r = 0.192, p < 0.01$) between $BPTA_{(0.5, 1.0, 2.0 \text{ kHz})}$ and the HINT composite score. The HINT composite score is the average of the thresholds across listening conditions with and without the spatial separation of the target speech and masking noise. Because the differences in results cannot be attributed to differences in protocols, they must be due to differences in participant characteristics. Specifically, the audibility of the test stimuli may have contributed to the differences in outcomes between studies.

Nilsson et al (1994) demonstrated the effect of decreased audibility of the test stimuli on speech recognition in noise ability using the HINT. They investigated the effect of decreased signal and noise bandwidth on the speech recognition in noise thresholds and found that as bandwidth decreased, the HINT thresholds became more positive (poorer). Nilsson and colleagues noted that patients with elevated pure-tone thresholds may experience a decrease in stimulus audibility when compared with participants with normal pure-tone thresholds. In other words, a hearing loss may result in partial audibility of the target speech and noise stimuli.

Purpose

The purpose of the present study was to determine the effect of stimulus audibility on the relationship between PTA and speech recognition in noise ability. For this retrospective study, participants were categorized

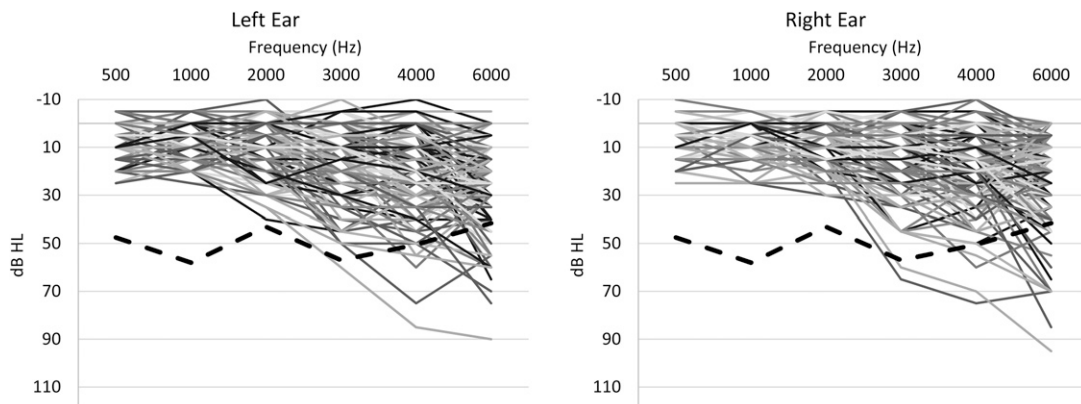


Figure 2. Audiograms for all participants from data set 1 ($n = 215$). The black dashed line represents the average level of the 65 dB(A) HINT stimuli through a TDH-39 earphone.

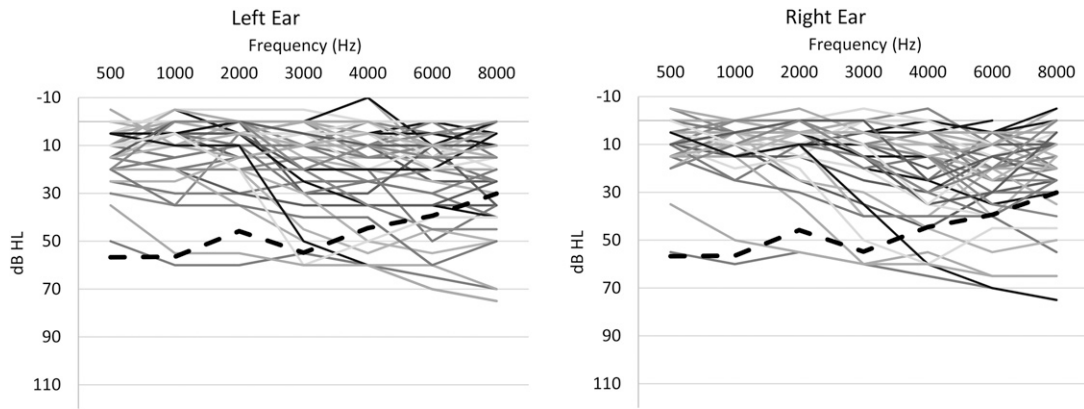


Figure 3. Audiograms for all participants from data set 2 (n = 55). The black dashed line represents the average level of 65 dB(A) HINT stimuli through a Sennheiser earphone. Seventeen of the participants in data set 2 were tested with the TDH-39 headphones.

as having either “complete” or “partial” audibility of the target speech and noise delivered at 65 dBA. It was hypothesized that when comparing data sets, a greater proportion of participants with partial audibility of the target speech and noise stimuli would be associated with a stronger relationship between BPTA and speech recognition in noise ability. The results of this study have potential implications for the conventional use of PTA for the determination of the ability to perceive speech in noisy environments as described by the AMA (2008) and the WHO (Mathers et al, 2000; Stevens et al, 2013).

METHODS

Permission to use the study data was obtained from the internal review board at St. Vincent Medical Center in Los Angeles, California. HINT and pure-tone threshold data were obtained from two different studies. Data set 1 (n = 215) was taken from Vermiglio et al (2012) and data set 2 (n = 55) was taken from previously unpublished data from Vermiglio and Soli (2008).

HINT Protocol

HINT thresholds for both data sets had been obtained with the standard HINT protocol. The target speech and steady-state speech-shaped noise stimuli were presented under headphones in a simulated soundfield environment using Knowles Electronics Mannequin for Auditory Research head-related transfer function. The steady-state speech-shaped noise was presented at a fixed level of 65 dBA. Complete details for this protocol are presented by Vermiglio et al (2012). As

noted previously, the results from data set 2 revealed a relatively strong relationship and the results from data set 1 revealed a relatively weak relationship between $BPTA_{(0.5, 1.0, 2.0 \text{ kHz})}$ and the HINT composite score. Although the relationship found for data set 1 is statistically significant, it is too weak to be used to predict the ability to recognize speech in noise from $BPTA_{(0.5, 1.0, 2.0 \text{ kHz})}$.

Pure-Tone Threshold Data

The pure-tone thresholds for data set 1 were measured for audiometric frequencies 250, 500, 1000, 2000, 3000, 4000, and 6000 Hz. The pure-tone thresholds for data set 2 were obtained for 500, 1000, 2000, 3000, 4000, 6000, and 8000 Hz. The pure-tone thresholds for data sets 1 and 2 are presented in Figures 2 and 3, respectively. According to the inclusion criteria from Vermiglio et al (2012), participants with normal pure-tone thresholds (≤ 25 dB HL, 250–6000 Hz) and participants with elevated pure-tone thresholds (> 25 dB HL, 2000–6000 Hz) were included in the study for data set 1 (Figure 2). The pure-tone thresholds for data set 2 included participants with normal (≤ 25 dB HL) or elevated (> 25 dB HL) pure-tone thresholds (500–8000 Hz). The black dashed lines in Figures 2 and 3 represent the average level of the 65 dBA HINT stimuli (target speech and steady-state speech-shaped noise).

Creation of Groups with Complete and Partial Audibility of the HINT Stimuli

For the present study, only the data from the Noise Front condition were selected for analysis. Recall that

Table 1. Average Third-octave Levels of 65 dB(A) HINT Stimuli in dB HL

Headphone	500 Hz	1000 Hz	2000 Hz	3150 Hz	4000 Hz	6300 Hz	8000 Hz
TDH-39	47.5	58.0	43.0	57.0	50.5	41.5	27
Sennheiser	56.7	56.5	45.8	54.8	44.5	39.4	30.1

Table 2. Descriptive Statistics for the BPTAs for Each Data Set

Variable	Data Set	n	Mean (dB HL)	Standard Deviation	Maximum (dB HL)	Minimum (dB HL)	Range (dB)
BPTA _(0.5, 1.0, 2.0 kHz)	1	215	6.20	5.20	25.00	-4.17	29.17
BPTA _(0.5, 1.0, 2.0, 3.0 kHz)	1	215	7.70	6.03	28.13	-1.88	30.00
BPTA _(0.5, 1.0, 2.0, 3.0, 4.0 kHz)	1	215	9.39	7.01	36.00	-2.00	38.00
BPTA _(0.5, 1.0, 2.0, 4.0 kHz)	1	215	8.69	6.28	30.00	-3.13	33.13
BPTA _(0.5, 1.0, 2.0, 3.0, 4.0, 6.0 kHz)	1	215	11.26	7.51	45.42	-0.83	46.25
BPTA _(0.5, 1.0, 2.0 kHz)	2	55	9.91	9.99	56.67	-0.83	57.50
BPTA _(0.5, 1.0, 2.0, 3.0 kHz)	2	55	10.99	10.84	56.88	-1.25	58.13
BPTA _(0.5, 1.0, 2.0, 3.0, 4.0 kHz)	2	55	12.32	11.86	58.00	-1.00	59.00
BPTA _(0.5, 1.0, 2.0, 4.0 kHz)	2	55	11.84	11.12	58.13	-0.63	58.75
BPTA _(0.5, 1.0, 2.0, 3.0, 4.0, 6.0 kHz)	2	55	13.79	12.56	59.58	0.42	59.17

for both data sets, the steady-state speech-shaped noise was presented at 65 dBA. This noise has the same long-term average spectrum as the target speech materials (Nilsson et al, 1994). Complete and partial HINT stimuli audibility groups were created to evaluate the contribution of stimulus audibility to the relationship between BPTA and speech recognition in noise thresholds (HINT Noise Front condition). The level of the steady-state speech-shaped noise (and consequently the average long-term spectrum of the target speech stimuli) was measured at third-octave bands for the TDH-39 supra-aural earphones (Telephonics Corporation, Farmingdale, NY) used for data set 1 and 17 of the 55 participants in data set 2. The same procedure was conducted for the Sennheiser circumaural earphones used for the remaining thirty-eight participants in data set 2. These levels were obtained in dB SPL. The SPL to HL conversion factors from ANSI S3.6-1996 were used to convert the third octave band levels to dB HL and are presented in Table 1. ANSI S3.6-1996 was the standard used during the collection of the pure-tone thresholds for both data sets. The third-octave band levels in Table 1 have been plotted for the TDH-39 earphones in Figure 2 and the Sennheiser earphones in Figure 3. Participants with pure-tone thresholds totaling at least 8 dB above the level of the HINT stimuli delivered at 65 dBA were assigned to the partial audibility group for each data set. All other

participants were assigned to the complete audibility group for each data set.

Pearson correlation coefficients were calculated for the HINT Noise Front thresholds versus BPTA. The BPTA was used for comparison with the HINT Noise Front thresholds because the HINT is a binaural test of the ability to recognize speech in noise. Five different BPTA were calculated: BPTA_(0.5, 1.0, 2.0 kHz), BPTA_(0.5, 1.0, 2.0, 3.0 kHz), BPTA_(0.5, 1.0, 2.0, 3.0, 4.0 kHz), BPTA_(0.5, 1.0, 2.0, 4.0 kHz), and BPTA_(0.5, 1.0, 2.0, 3.0, 4.0, 6.0 kHz). Recall that a previous method for the determination of hearing impairment used by the AMA (Lierle, 1959) used a monaural PTA_(0.5, 1.0, 2.0 kHz). The current AMA method for the determination of hearing impairment (AMA, 2008) uses a monaural PTA_(0.5, 1.0, 2.0, 3.0 kHz), and the WHO method (Mathers et al, 2000) uses monaural PTA_(0.5, 1.0, 2.0, 3.0, 4.0 kHz).

Statistical Analyses

All statistical analyses were performed in JMP Pro 12 (SAS Institute Inc., Cary, NC). Pearson correlation coefficients were obtained to quantify the relationships between BPTA and HINT Noise Front thresholds. Regression analyses were conducted to evaluate the contribution of age for these relationships. A two-way ANOVA was conducted to investigate the main effects for data set and stimulus audibility, and the interaction on speech recognition

Table 3. Descriptive Statistics for the Noise Front Thresholds for Each Data Configuration

Variable	Data	n	Mean (dB SNR)	Standard Deviation	Maximum (dB SNR)	Minimum (dB SNR)	Range (dB)
Noise Front	Sets 1 & 2	270	-2.49	1.15	4.50	-5.30	9.80
Noise Front	Set 1	215	-2.68	0.98	1.27	-5.30	6.57
Noise Front	Set 2	55	-1.78	1.47	4.50	-3.60	8.10
Noise Front	Set 1, complete	194	-2.69	0.97	0.00	-5.30	5.30
Noise Front	Set 1, partial	21	-2.51	1.10	1.27	-4.19	5.46
Noise Front	Set 2, complete	46	-2.09	0.83	0.50	-3.60	4.10
Noise Front	Set 2, partial	9	-0.19	2.69	4.50	-3.40	7.90

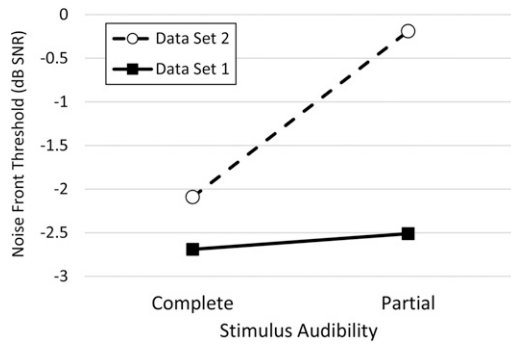


Figure 4. Interaction of data set and stimulus audibility.

in noise ability. A significance level of 0.05 was adopted for all statistical tests.

RESULTS

Descriptive Statistics

The descriptive statistics for BPTA for data sets 1 and 2 are presented in Table 2. Generally, as the high-frequency contribution to BPTA increased, the mean BPTA increased. The range of BPTA was also greater for data set 2 than data set 1. This is a reflection of the greater proportion of participants with elevated pure-tone thresholds for data set 2 than data set 1.

Table 3 shows the descriptive statistics for the Noise Front thresholds for various configurations of the data sets and complete and partial audibility groups. The more negative thresholds indicate better performances. A two-way ANOVA was conducted to investigate the effects of data set and stimulus audibility on Noise Front thresholds. This analysis revealed that both main effects and their interaction were statistically significant

(*F* values are 41.6421 for data set, 21.2270 for stimulus audibility, and 14.3986 for interaction; all *p*-values < 0.05). The interaction is plotted in Figure 4. A greater difference was found between the data sets for the Noise Front thresholds obtained with partial rather than complete audibility of the stimuli.

Relationship between PTA and Speech Recognition in Noise

For illustration, the relationship between $BPTA_{(0.5, 1.0, 2.0, 3.0, 4.0, 6.0 \text{ kHz})}$ versus HINT Noise Front thresholds for each data set is presented in Figure 5. A statistically significant but relatively weak relationship was found between variables for data set 1 ($r = 0.166, p = 0.015$) and a relatively strong relationship between variables was found for data set 2 ($r = 0.692, p < 0.001$). Table 4 shows the Pearson correlation coefficients for the Noise Front thresholds versus $BPTA_{(0.5, 1.0, 2.0, 3.0, 4.0, 6.0 \text{ kHz})}$ for various group configurations. Weak and nonsignificant relationships were found between these variables for the complete audibility groups within each data set. Strong and statistically significant relationships were found between variables for the partial audibility groups within each data set ($p < 0.05$). Note that the proportion of participants with partial audibility was larger for data set 2 (16%) than for data set 1 (10%).

The effect of stimulus audibility on the relationship between BPTA and speech recognition in noise ability is clearly illustrated in Figure 6. A weak and nonsignificant relationship was found between $BPTA_{(0.5, 1.0, 2.0, 3.0, 4.0, 6.0 \text{ kHz})}$ versus HINT Noise Front thresholds for all of the data with complete audibility combined across data sets 1 and 2 ($r = 0.060, p = 0.356$) and a strong relationship was found between these

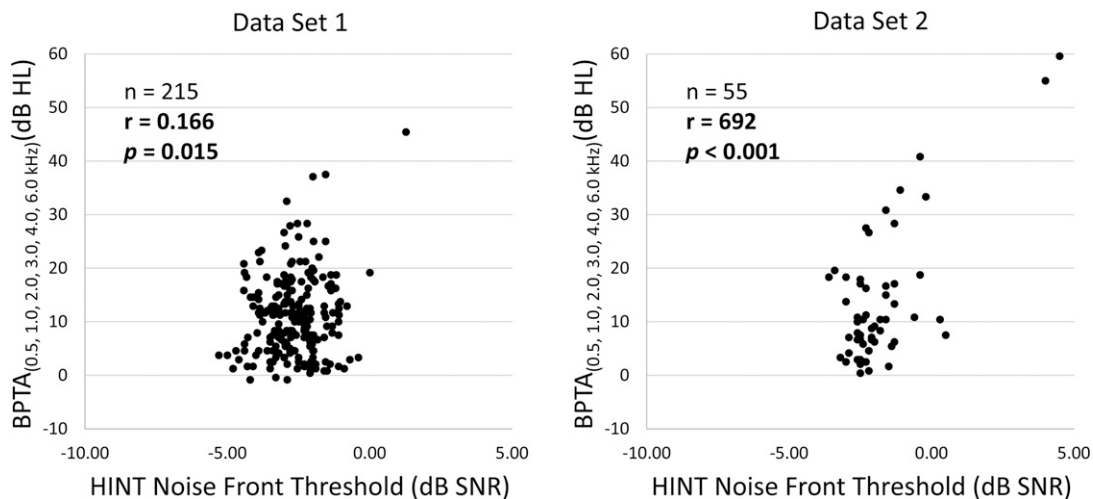


Figure 5. Scatter plots of the HINT Noise Front thresholds for data set 1 (left) and data set 2 (right) plotted against $BPTA_{(0.5, 1.0, 2.0, 3.0, 4.0, 6.0 \text{ kHz})}$. The statistically significant correlations are presented in bold font.

Table 4. Pearson Correlation Coefficient (r) for BPTA_(0.5, 1.0, 2.0, 3.0, 4.0, 6.0 kHz) Versus the HINT Noise Front Thresholds

Data Set	Audibility Group	n	Percent of Total Participants	r	p
1	All data	215	100%	0.166	0.015
	Complete	194	90%	0.083	0.251
	Partial	21	10%	0.586	0.005
2	All data	55	100%	0.692	< 0.001
	Complete	46	84%	0.124	0.414
	Partial	9	16%	0.961	< 0.001

Note: Statistically significant relationships are in bold font.

variables for all partial audibility data combined across data sets ($r = 0.863, p < 0.001$). It appears that the proportion of partial audibility data is driving the strength of the relationships between BPTA and HINT Noise Front thresholds across all data for data sets 1 and 2.

Table 5 shows the correlation coefficients for the relationships between five different BPTAs versus the HINT Noise Front thresholds for all data (n = 270), all complete audibility data (n = 240), all partial audibility data (n = 30), and age for data sets 1 and 2 combined. The same relationship is shown between the monaural PTA_(0.5, 1.0, 2.0, 4.0 kHz) for the better ear as used in the WHO method for the determination of hearing impairment. Moderate and statistically significant correlations were found between PTA and Noise Front thresholds for all data combined ($p < 0.01$). Weak and nonsignificant relationships were found for all complete audibility data. Strong and statistically significant correlations were found for all partial audibility data ($p < 0.01$). In addition, modest and statistically significant correlations were found between age and each of the PTAs ($p < 0.01$).

Multiple regression models were used to investigate the associations between the Noise Front thresholds and BPTA_(0.5, 1.0, 2.0, 3.0, 4.0, 6.0 kHz) while controlling for age. For all participants combined (n = 270) after controlling for age, the Noise Front thresholds were still significantly associated with BPTA_(0.5, 1.0, 2.0, 3.0, 4.0, 6.0 kHz)

($p < 0.001$). For all participants with complete audibility of the stimulus (n = 240) after controlling for age, the Noise Front thresholds were not significantly associated with BPTA_(0.5, 1.0, 2.0, 3.0, 4.0, 6.0 kHz) ($p = 0.541$). For all participants with partial audibility of the stimulus (n = 30) after controlling for age, the Noise Front thresholds were still significantly associated with BPTA_(0.5, 1.0, 2.0, 3.0, 4.0, 6.0 kHz) ($p < 0.001$).

DISCUSSION

Results of this retrospective study indicate that audibility of the test stimuli is a factor that may affect the relationship between speech recognition in noise thresholds versus PTA. The effect of stimulus audibility may explain the strong relationship reported by Tschopp and Züst (1994) where all 33 participants had a hearing impairment and of these and 19 had a profound high-frequency hearing loss. The relatively strong relationship found between PTA and speech recognition in noise ability in this study was most likely influenced by a relatively large portion of the participants with partial audibility of the stimuli. The influence of stimulus audibility can also be seen across the studies shown in Figure 1. Studies with the highest stimulus levels (Wilson, McArdle, et al, 2007) also had some of the lowest correlation coefficients for the relationship between PTA and speech recognition in noise performance. This was most likely due to the overall improvement in

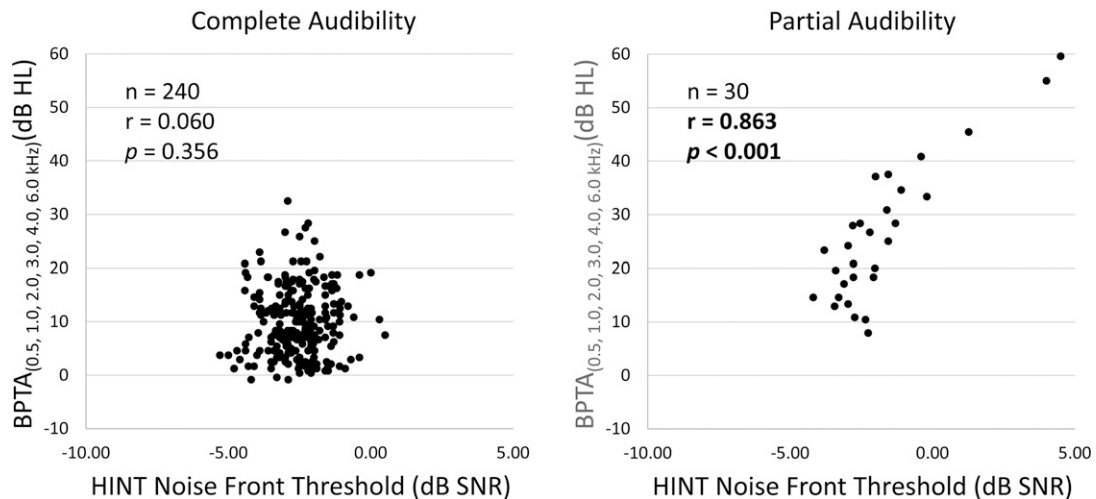


Figure 6. Scatter plots of the HINT Noise Front thresholds for the complete audibility group (left) and the partial audibility group (right) plotted against BPTA_(0.5, 1.0, 2.0, 3.0, 4.0, 6.0 kHz). The statistically significant correlation is presented in bold font.

Table 5. Pearson Correlation Coefficients (*r*) for HINT Noise Front Thresholds Versus PTA and Versus Age

PTA	Noise Front Threshold, All Data (n = 270)	Noise Front Threshold, All Complete Audibility Data (n = 249)	Noise Front Threshold, All Partial Audibility Data (n = 21)	Age, All Data (n = 270)
BPTA _(0.5, 1.0, 2.0 kHz)	0.409 (< 0.001)	0.084 (0.196)	0.862 (< 0.001)	0.300 (< 0.001)
BPTA _(0.5, 1.0, 2.0, 3.0 kHz)	0.402 (< 0.001)	0.060 (0.356)	0.876 (< 0.001)	0.335 (< 0.001)
BPTA _(0.5, 1.0, 2.0, 3.0, 4.0 kHz)	0.401 (< 0.001)	0.079 (0.225)	0.851 (< 0.001)	0.355 (< 0.001)
BPTA _(0.5, 1.0, 2.0, 4.0 kHz)	0.424 (< 0.001)	0.104 (0.107)	0.885 (< 0.001)	0.355 (< 0.001)
BPTA _(0.5, 1.0, 2.0, 3.0, 4.0, 6.0 kHz)	0.385 (< 0.001)	0.060 (0.356)	0.863 (< 0.001)	0.360 (< 0.001)
WHO monaural PTA _(0.5, 1.0, 2.0, 4.0 kHz) for the better ear	0.411 (< 0.001)	0.082 (0.208)	0.859 (< 0.001)	0.352 (< 0.001)

Notes: Probabilities (*p*) are in parentheses. Statistically significant relationships are in bold font.

stimulus audibility when compared with the studies with lower noise levels.

Wilson, McArdle, et al (2007) measured speech recognition in noise ability for a group of participants with normal pure-tone thresholds and a group with hearing loss. As shown in Figure 1, they used four different speech recognition in noise tests; the BKB-SIN, the QuickSIN, WIN, and the HINT. Results of this study demonstrated that the strongest relationship between PTA_(1.0, 2.0, 3.0, 4.0 kHz) and speech recognition in noise ability was found with the WIN ($r = 0.689$) and QuickSIN ($r = 0.557$) results and weaker relationships were found with the BKB-SIN ($r = 0.355$) and HINT results ($r = 0.323$). The true positive test results (sensitivity) were reported as follows: WIN, 99%; QuickSIN, 90%; BKB-SIN; 78%; and the HINT 72%. From these results, Wilson and colleagues concluded that “the QuickSIN and WIN materials are more sensitive measures of recognition performance in background noise than are the BKB-SIN and HINT materials.” However, Wilson and colleagues did not evaluate the sensitivity of the speech recognition in noise tests for the detection of speech recognition in noise disorder. Instead, they determined the sensitivity of these tests for the detection of elevated PTAs. Stated differently, Wilson et al used the pure-tone threshold test as the “gold” or reference standard for the evaluation of the sensitivity of the speech recognition in noise tests. This is known because the non-disordered and disordered groups were created based on pure-tone threshold results where PTAs greater than 20 dB HL represented a hearing impairment. Wilson, McArdle, et al (2007) made a decision on the suitability of a speech recognition in noise test based on the strength of the relationship between speech recognition in noise ability and PTA. In contrast, Wilson, Carnell, et al (2007) wrote that “there is an abundance of data that indicates speech-recognition performance in background noise cannot be predicted with any degree of certainty by either pure-tone thresholds or by speech-recognition performance in quiet.”

Conclusions

The relationship between PTA and speech recognition in noise ability may be influenced by the audibility

of the test stimuli. Specifically, the relationship between BPTA and HINT Noise Front thresholds cannot be predicted when the participants have complete audibility of the HINT stimuli. The proportion of participants with partial audibility of the speech and noise stimuli most likely influenced this relationship across the literature (Figure 1). This brings into question the convention of using PTA for the determination of the ability to perceive speech in the presence of noise as described in the AMA (AMA, 2008) and WHO (Mathers et al, 2000) methods for the determination of hearing impairment.

Evaluations of hearing impairment should continue to include pure-tone threshold testing. However, because PTA cannot be used to predict speech recognition in noise ability in conditions with complete audibility of the stimuli, a reliable, valid, and norm-referenced speech recognition in noise test should also be incorporated in any evaluation of the ability to hear. The speech recognition in noise test should approximate “real-world” target speech and noise listening conditions.

Study Limitations and Future Research

The effect of stimulus audibility on the relationship between PTA and speech recognition in noise results using the HINT may not apply to other types of stimuli and test protocols. Future studies should investigate the effect of stimulus audibility with alternative test parameters. Future work should also include both younger and older participants.

Acknowledgments. The first author would like to thank Brenda Vermiglio, MA, for her helpful comments.

REFERENCES

- AAO-ACO. (1979) Guide for the evaluation of hearing handicap. *JAMA* 241:2055–2059.
- AMA. (2008). In: Rondinelli RD ed. *Guides to the Evaluation of Permanent Impairment*. 6th ed. Chicago, IL: American Medical Association.

- Demorest ME, Erdman SA. (1987) Development of the communication profile for the hearing impaired. *J Speech Hear Disord* 52(2):129–143.
- Dobie RA. (2011) The AMA method of estimation of hearing disability: a validation study. *Ear Hear* 32(6):732–740.
- Killion MC, Niquette PA, Gudmundsen GI, Revit LJ, 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(4 Pt 1):2395–2405.
- Lierle DM. (1959) (For the subcommittee on noise of the AAOO committee on conservation of hearing). Guide for the evaluation of hearing impairment. *Trans Am Acad Ophthalmol Otolaryngol* 63:236–238.
- Lyregaard PE. (1982) Frequency selectivity and speech intelligibility in noise. *Scand Audiol Suppl* 15:113–122.
- Mathers C, Smith A, Concha M. (2000) Global burden of hearing loss in the year 2000. *Glob Burden Dis* 18:1–30.
- Nilsson M, Soli SD, Sullivan JA. (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(2):1085–1099.
- Niquette P, Arcaroli J, Revit L, Parkinson A, Staller S, Skinner M, Killion M. (2003) Development of the BKB-SIN Test. Paper presented at the American Auditory Society, Scottsdale, AZ.
- Plomp R, Mimpen AM. (1979) Improving the reliability of testing the speech reception threshold for sentences. *Audiology* 18(1):43–52.
- Saunders GH, Forsline A, Fausti SA. (2004) The performance-perceptual test and its relationship to unaided reported handicap. *Ear Hear* 25(2):117–126.
- Smits C, Kapteyn TS, Houtgast T. (2004) Development and validation of an automatic speech-in-noise screening test by telephone. *Int J Audiol* 43(1):15–28.
- Smoorenburg GF. (1992) Speech reception in quiet and in noisy conditions by individuals with noise-induced hearing loss in relation to their tone audiogram. *J Acoust Soc Am* 91(1):421–437.
- Stevens G, Flaxman S, Brunskill E, Mascarenhas M, Mathers CD, Finucane M; Global Burden of Disease Hearing Loss Expert Group. (2013) Global and regional hearing impairment prevalence: an analysis of 42 studies in 29 countries. *Eur J Public Health* 23(1):146–152.
- Tschopp K, Zust H. (1994) Performance of normally hearing and hearing-impaired listeners using a German version of the SPIN test. *Scand Audiol* 23(4):241–247.
- Vermiglio AJ. (2007) *Speech Recognition in Noise*. (AuD Doctoral Project), Central Michigan University.
- Vermiglio AJ. (2008) The American English hearing in noise test. *Int J Audiol* 47(6):386–387.
- Vermiglio AJ, Soli S. (2008) Is there a relationship between pure tone thresholds and speech recognition in noise ability? Research Poster Presented at the International Hearing Aid Research Conference (IHCON). Lake Tahoe, CA.
- Vermiglio AJ, Soli SD, Freed DJ, Fisher LM. (2012) The relationship between high-frequency pure-tone hearing loss, hearing in noise test (HINT) thresholds, and the articulation index. *J Am Acad Audiol* 23(10):779–788.
- Wilson RH, Carnell CS, Cleghorn AL. (2007) The words-in-noise (WIN) test with multitalker babble and speech-spectrum noise maskers. *J Am Acad Audiol* 18(6):522–529.
- Wilson RH, McArdle RA, Smith SL. (2007) An evaluation of the BKB-SIN, HINT, QuickSIN, and WIN materials on listeners with normal hearing and listeners with hearing loss. *J Speech Lang Hear Res* 50(4):844–856.
- Zust H, Tschopp K. (1993) Influence of context on speech understanding ability using German sentence test materials. *Scand Audiol* 22(4):251–255.