Korean J Otorhinolaryngol-Head Neck Surg 2016;59(10):725-9 / pISSN 2092-5859 / eISSN 2092-6529
http://dx.doi.org/10.3342/kjorl-hns.2016.59.10.725

The Best-Matched Pure Tone Average and Speech Recognition Threshold for Different Audiometric Configurations

Jeong Min Kim, Mi Sun Na, Ki Hwan Jung, Soo Hyeong Lee, Jae Sang Han, Oh Hyeong Lee, and So Young Park
Department of Otorhinolaryngology-Head and Neck Surgery, College of Medicine, The Catholic University of Korea, Seoul, Korea

순음청력도 유형별로 어음인지역치와 가장 일치하는 순음역치 평균법

김정민 · 나미선 · 정기환 · 이수형 · 한재상 · 이오형 · 박소영
가톨릭대학교 의과대학 이비인후-두경부외과학교실

Received May 23, 2016
Revised July 26, 2016
Accepted August 29, 2016
Address for correspondence
So Young Park, MD, PhD
Department of Otorhinolaryngology-Head and Neck Surgery,
Yeouido St. Mary's Hospital,
College of Medicine,
The Catholic University of Korea,
10 63-ro, Yeongdeungpo-gu,
Seoul 07345, Korea
Tel +82-2-3779-1044
Fax +82-2-786-1149
E-mail sypak@catholic.ac.kr

Background and Objectives The agreement between pure-tone average (PTA) and speech recognition threshold (SRT) has become more important with the increasing demands for medical certification. The purpose of this study was to explore the relationships between the SRT and several variations of PTA, and to determine which PTA formula would provide the best agreement with SRT for different audiometric configurations.
Subjects and Method Audiological data on 783 ears were retrospectively collected. The air-conduction PTAs were calculated using five different formulas: three-frequency average (3FA), weighted three-frequency average (W3FA), four-frequency average (4FA), weighted four-frequency average (W4FA), and six-frequency average (6FA). The audiometric configuration was classified into five categories. The PTA-SRT relationships were analyzed using correlation and simple linear regression for each audiometric configuration.
Results Highest correlation was observed between the SRT and W3FA for all audiometric configurations with the correlation coefficient of 0.964 as a whole. The SRT and 3FA were best-matched in the linear regression models for overall/flat/high frequency gently sloping/low frequency ascending; the SRT and W3FA were best-matched for high frequency steeply sloping (HFSS); the SRT and 4FA were best-matched for miscellaneous audiograms.
Conclusion The most stable PTA variations that make the best-matched pairs with SRT for any audiogram are the conventional 3FA and W3FA doubling 1 kHz threshold. The addition of frequencies higher than 2 kHz to a PTA formula seems to have impeded the PTA-SRT agreement, especially for HFSS audiograms. W3FA should be the method of choice in predicting SRT from PTA for HFSS audiograms.
Korean J Otorhinolaryngol-Head Neck Surg 2016;59(10):725-9

Key Words Audiogram · Functional hearing loss · Medical certificate · Pure-tone threshold · Three-frequency average.

Introduction

A pure tone audiogram provides the basic information on hearing acuity, and the pure-tone average (PTA) with speech intelligibility represents the total hearing status of a person. Speech recognition threshold (SRT) is not only the basis for speech discrimination testing, but also indicates the reliability of a pure tone audiogram. The consistency between the PTA and SRT would be important when the patient is likely to have a functional hearing loss (FHL). Although the primary goal of obtaining and reporting PTA exists in quantifying the patient's hearing level most appropriately but not in predicting SRT, a discrepancy between the PTA and SRT has been considered as one of the essential criteria for FHL. SRT has been documented to be frequently lower than PTA by 12 dB or more in subjects with FHL. The hypothesized reasons are based on three facts: the FHL patients use a comfortable loudness reference to make response decisions to suprathreshold stimuli; tones and spondees sound equal in loudness at equal sound pressure levels (SPLs); behavioral thresholds are measured in dB hearing level (HL) and greater SPL is required to calibrate 0 dB HL for spondee threshold than for tonal thresholds. Therefore, a PTA-SRT discrepancy does not occur in patients who are responding at true thresholds.[1] The attempts to predict SRT from pure-tone thresholds (PTTs) have been made in two ways from the initial period of audiometry.[2,3] Some investigators have suggested the methods of simply averaging the selected PTTs, mostly at frequencies 0.5–2 kHz having high acoustic energy for speech.[4,5] Others proposed the sum of PTTs weighted differentially using multiple regression equations.[6-8] However, the use of a multiple regression equation, the practical applicability of which may be in doubt, does not seem to improve the predictions obtainable by the PTAs.[3] Since the PTA has been defined by ISO/1964 and ANSI/1969 as a mean air-conduction threshold level for usual conversational frequencies, various formulas have often been employed to calculate PTAs for clinical or medicolegal purposes. Conventionally, the severity of hearing loss has been classified according to a widely used system by ANSI/1969, i.e., the three-frequency average method (0.5, 1, and 2 kHz); the American Academy of Otolaryngology-Head and Neck Surgery (AAO-HNS) standards for reporting hearing loss include PTA at 0.5, 1, 2, and 3 kHz[9]; PTAs weighting 1 kHz (0.5, 1, 1, and 2 kHz), or 1 and 2 kHz (0.5, 1, 1, 2, 2, and 4 kHz) have been adopted for medical certificates in our country; a full range average across the test frequencies could give a concise information for an audiogram.[5]

The average method would be simpler and more applicable than the multiple regression method in prediction of SRT from PTTs in otology clinics. Before application, the pattern of pure tone loss should be taken into account first. Although Carhart and Porter[10] reported the influence of audiogram patterns on multiple regression equations using frequencies through 0.5–2 kHz, there have been few studies on the effect of audiometric patterns on the PTA-SRT relationship, especially using various versions of the PTA method. At this point, it would be worth restudying the methods of predicting SRT from PTA in conjunction with rapidly increasing demands for medicolegal certification over recent years. The purpose of this study was to explore the relationships between the SRT and several variations on PTAs commonly used in clinical practice, and to determine which variation on PTA would provide the best agreement with SRT for different audiometric configurations.

Subjects and Method

Subjects and audiometry

Audiological data were retrospectively collected on patients who underwent audiometry in our otology clinic between Apr and Oct in 2015. The ears with normal tympanogram and word recognition score $\geq$70% were included in the analysis except those with PTTs $\leq$15 dB or $>$90 dB hearing level (HL) at all test frequencies or those with a possible FHL. PTTs were acquired between 125 and 8000 Hz using a well-standardized audiometric procedure (GSI 61; Grason-Stadler Inc., Eden Prairie, MN, USA). Standard supra-aural earphones, the Telephonics TDH-50 were applied. The air-conduction PTAs were calculated using five different formulas and rounded to the nearest whole number: conventional three-frequency average (3FA)=(0.5 kHz+1 kHz+2 kHz)/3; weighted three-frequency average (W3FA)=(0.5 kHz+1 kHz+1 kHz+2 kHz)/4; four-frequency average (4FA)=(0.5 kHz+1 kHz+2 kHz+3 kHz)/4; weighted four-frequency average (W4FA)=(0.5 kHz+1 kHz+1 kHz+2 kHz+2 kH+4 kHz)/6; six-frequency average (6FA)=(0.5 kHz+1 kHz+2 kHz+3 kHz+4 kH+6 kHz)/6. Using the same equipment, SRT was obtained with monitored live-voice presentations of Hahm's spondee word lists.[11,12] Word recognition was scored in terms of the percentage of phonetically balanced words that were correctly identified from Hahm's monosyllabic word list. All tests were administered monoaurally.

Classification of audiometric configurations

To characterize the audiometric configuration, pure tone audiograms were classified into seven categories modified from the previous reports.[13,14] A Flat audiogram was defined as the one where the thresholds across all frequencies did not vary more than 15 dB HL from each other, or where the differences between the mean threshold of 250 and 500 Hz, that of 1 and 2 kHz, and that of 4 and 8 kHz were $<$15 dB HL; a high frequency gently sloping (HFGS) audiogram, where the difference between the mean threshold of 0.5 and 1 kHz and that of 4 and 8 kHz was ≥15 dB and <30 dB; a high frequency steeply sloping (HFSS) audiogram, where the difference between the mean threshold of 0.5 and 1 kHz and that of 4 and 8 kHz was ≥30 dB; a low frequency ascending (LFA) audiogram, where the mean threshold of 250 and 500 Hz was ≥15 dB poorer than that of 4 and 8 kHz; a mid-frequency U-shape (MFU) audiogram, where the mean threshold of the mid-frequencies was ≥15 dB poorer than those of the lower and higher frequencies; a mid-frequency reverse U-shape (MFRU), where the mean threshold of the mid-frequencies was ≥15 dB better than those of the lower and higher frequencies. The other shapes were classified as atypical audiograms. However, the numbers of MFU and MFRU were too small for statistics, the last three configurations were taken together as miscellaneous audiograms.

Statistical analysis

Statistical analysis was performed using SPSS 18.0 software (SPSS Inc., Chicago, IL, USA) and a two-tailed p value less than 0.05 was considered to be significant. The relationships between the SRT and five variations on PTA were analyzed in three ways: correlation, simple linear regression, and mean comparison. Pearson's correlation coefficients (r) were obtained between the SRT and PTAs overall and within five groups of audiometric configurations. Simple linear regression analysis was done to define a more definite relationship between the SRT (y) and PTAs (x) overall and within groups. The regression coefficient and standard error of the estimate (SEE, σ_{est}) were obtained for every PTA-SRT match. Mean PTA-SRT difference (PTA minus SRT) was analyzed using paired t test for overall, flat, HFGS, and HFSS audiograms; Wilcoxon signed-rank test for LFA and miscellaneous audiograms.

Results

Subjects were 783 ears (382 right and 401 left) from 460 patients (186 men and 274 women aged mean 56.6 years in the range of 7−86 years). Flat audiograms were demonstrated in 313 (40.0%) ears; HFGS in 228 (29.1%); HFSS in 168 (21.5%); LFA in 34 (4.3%); miscellaneous in 40 (5.1%). The SRT ranged from 5 to 85 dB HL. Table 1 shows the detailed data of the correlation coefficients, regression coefficients and constants, and SEEs between the SRT and five variations on PTA overall and for each audiometric configuration. 1) The correlation coefficients (r) were >0.9 (0.915−0.976) in all matches except SRT vs 6FA as a whole and except SRT vs 4FA/6FA for HFSS.

Table 1. The relationships between speech recognition threshold (SRT) and various pure-tone averages (PTAs) for different audiometric configurations

SRT vs	r	β	α	σ_{est}
All (n=783)				
3FA	0.958	0.950*	2.126	4.5
W3FA	0.964*	0.944	2.037	4.3*
4FA	0.939	0.918	2.039	5.5
W4FA	0.945	0.908	1.047	5.2
6FA	0.890	0.814	1.031	7.3
Flat (n=313)				
3FA	0.951	0.933*	2.260	4.6
W3FA	0.954*	0.920	2.107	4.4*
4FA	0.940	0.927	2.731	5.1
W4FA	0.949	0.923	1.975	4.7
6FA	0.939	0.925	1.310	5.1
HFGS (n=228)				
3FA	0.969	0.958*	2.093	4.2
W3FA	0.975*	0.956	1.968	3.8*
4FA	0.960	0.948	0.907	4.7
W4FA	0.970	0.944	−0.394	4.1
6FA	0.965	0.953	−4.948	4.4
HFSS (n=168)				
3FA	0.929	0.945	2.553	5.0
W3FA	0.942*	0.955*	2.386	4.6*
4FA	0.894	0.883	0.371	6.1
W4FA	0.915	0.928	−3.145	5.5
6FA	0.858	0.883	−9.086	7.0
LFA (n=34)				
3FA	0.959	0.976*	1.072	4.4
W3FA	0.968*	0.952	1.683	3.9*
4FA	0.954	1.099	1.623	4.7
W4FA	0.950	1.041	2.667	4.9
6FA	0.911	1.098	3.208	6.4
Misc. (n=40)				
3FA	0.968	0.967	1.103	4.6
W3FA	0.976*	0.935	1.818	4.0*
4FA	0.970	0.993*	1.630	4.5
W4FA	0.964	0.965	0.813	4.9
6FA	0.960	1.027	−2.801	5.1

Regression equation is SRT=α+β×PTA. All p values for r and β <0.001. The top coefficient closest to 1.0 and the minimal standard error of the estimate in each subset are marked with asterisks (*). r: correlation coefficient, β: regression coefficient (slope), α: constant (intercept), σ_{est}: standard error of the estimate (dB HL). 3FA: three-frequency average (0.5, 1, and 2 kHz), W3FA: weighted three-frequency average (0.5, 1, 1, and 2 kHz), 4FA: four-frequency average (0.5, 1, 2, and 3 kHz), W4FA: weighted four-frequency average (0.5, 1, 1, 2, 2, and 4 kHz), 6FA: six-frequency average (0.5, 1, 2, 3, 4, and 6 kHz), HFGS: high frequency gently sloping, HFSS: high frequency steeply sloping, LFA: low frequency ascending, Misc.: miscellaneous

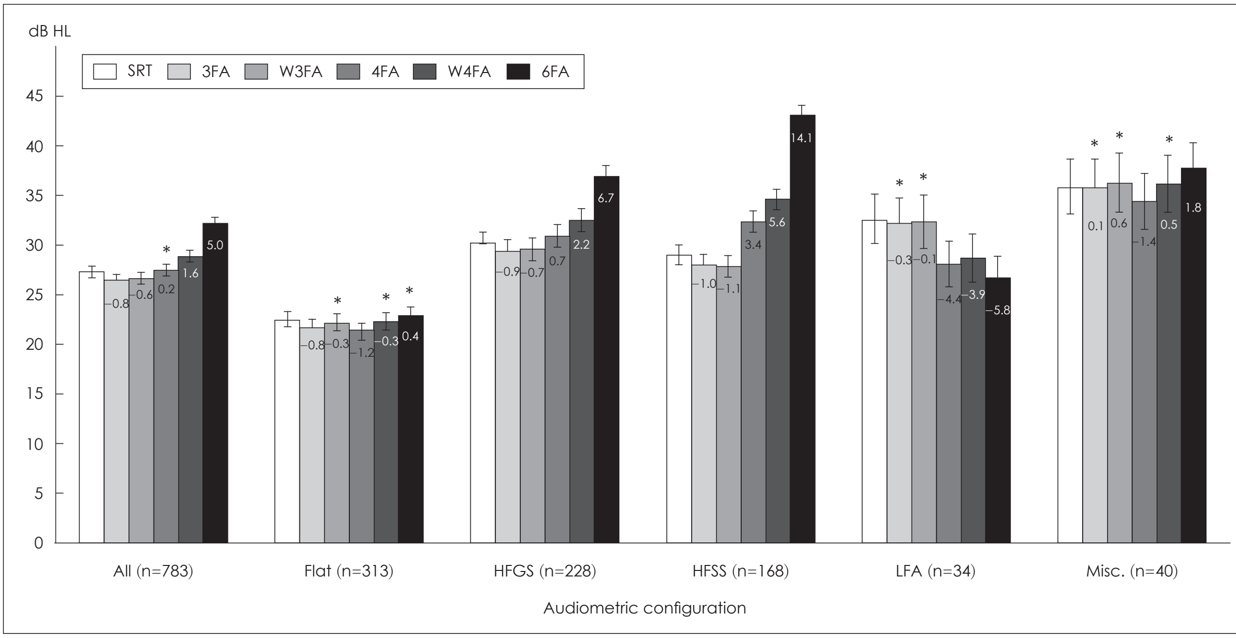

Fig. 1. Mean comparison of SRT and five variations on PTAs for different audiometric configurations. The mean difference (PTA minus SRT) is written in each PTA column. Error bars indicate standard error of mean. *$p \geq 0.05$ (no difference between SRT and PTA). Abbreviations as in Table 1. SRT: speech recognition threshold, PTA: pure-tone average, W3FA: weighted three-frequency average, 4FA: four-frequency average, W4FA: four-frequency average, 6FA: six-frequency average, HFGS: high frequency gently sloping, HFSS: high frequency steeply sloping, LFA: low frequency ascending, Misc.: miscellaneous.

2) All regression coefficients (β) approached 1.0 within the range of ±0.1 (0.908–1.099) except for the same matches as for the correlation coefficients. 3) The SEEs (σ_{est}) ranged from 3.8 to 7.3 dB HL. All p values for r and β were statistically significant (all ps <0.001). Fig. 1 demonstrates the mean comparison between the SRT and each variation on PTA for different audiometric configurations. The mean SRT-PTA differences were in the range of -5.8–14.1 dB HL as presented in Fig. 1.

Discussion

In the present study, the best-matched pair of the SRT and PTA variation was identified for each audiometric configuration primarily through correlation and simple linear regression analyses. The most common audiometric pattern was the flat audiogram followed by HFGS, HFSS, and LFA in order, which was consistent with the previous report.[10] The highest correlation was found between the SRT and W3FA for all audiometric configurations with a correlation coefficient of 0.964 as a whole. One point to be noted is that the correlation coefficient is only the measure of strength of a linearity regardless of the slope, which sometimes makes its clinical significance less important. In an extreme case of a horizontal linearity where the slope (β)=0, two variables are hardly considered to be related even if r=1, a perfect line. On the other hand, if β=1 in a linear regression model, one unit change of the x-value makes one unit change of the y-value. For another example, if β=1 and y-intercept (α)=0, the x-values become in complete agreement with the y-values. Most of the regression coefficients (β) in this study were very close to 1.0. The SRT and 3FA were best-matched in the linear regression models for overall/flat/HFGS/LFA audiograms (β=0.950/0.933/0/958/0/976); the SRT and W3FA were best-matched for HFSS (β=0.955); the SRT and 4FA were best-matched for miscellaneous audiograms (β=0.993). The SEEs (σ_{est}) were close to 5 dB for most of the PTA-SRT matches, less than 4 dB for two matches (SRT vs W3FA for HFGS/LFA), and more than 6 dB for four matches (SRT vs 4FA for HFSS and SRT vs 6FA for overall/HFSS/LFA). It is notable that minimal SEEs were found between the SRT and W3FA for all audiometric configurations with 4.3 dB as a whole. The SEE means that the SRT can be estimated correctly within $2\sigma_{est}$ from PTA in 95% of all cases. Our data showed that the actual SRT can be predicted within approximately ±10 dB error from any of the PTA variations in 95% of cases except for SRT vs 4FA/W4FA/6FA for HFSS and SRT vs 6FA for LFA.

In a previous literature, a correlation coefficient of 0.986 and SEE of 3.5 dB have been reported by the use of Fletcher's method (two-frequency average at two best thresholds at 0.5–

2 kHz); a correlation coefficient of 0.981 and SEE of 4.1 dB by the use of three-frequency average at 0.5–2 kHz,[3] which were very similar with our results for the pair of SRT and 3FA (r=0.958 and σ_{est}=4.5 when disregarding the configuration). Kim, et al.[15] have demonstrated the same correlation coefficient (r=0.938) with ours (r=0.939) between the PTA (4FA) and SRT (Hahm's list) in the hearing-impaired group. Coren and Hakstian[5] have reported correlation coefficients of 0.855–0.893, which are lower than ours, between the SRT and five subsets of PTAs without classification of audiogram patterns. Their combinations of tones were different from those of our study except for 3FA (r=0.893). The regression coefficient for 3FA (β=0.955) were almost identical with ours (0.950), but relatively large constant (α=19.116) compared to ours (2.126) resulted in PTA-SRT discrepancy over 15 dB.[5] With respect to the mean values, most of the PTAs were in line with SRT except 4FA, W4FA, and 6FA for sloping and ascending audiograms. The statistical significance of the mean difference between the SRT and PTA was less focused in this study because most of the difference were far less than 5 dB regardless of the statistical significance. It is reminded that SRT is measured in 5-dB steps. Kim and Lee[16] have reported that the differences between the PTA (3FA) and SRT (Hahm's list) were 2.4 dB in the conductive hearing loss, and 1.9 dB in the normal hearing groups. Our data showed much smaller differences between the SRT and 3FA with the range of 0.1–1.0 dB according to the audiogram patterns.

In summary, 1) for flat, HFGS, and LFA audiograms, 3FA showed the best agreement with SRT (β=0.933, 0.958, and 0.976 respectively) along with high correlations of $>$0.950 and small SEEs of $\leq$4.6 dB. 2) For HFSS audiograms, W3FA doubling 1 kHz threshold showed the best agreement with SRT (β=0.955) along with a correlation of 0.942 and an SEE of 4.6 dB. 3) For miscellaneous audiograms other than the four most common configurations, 4FA including 3 kHz threshold showed the best agreement with SRT (β=0.993) along with a correlation of 0.970 and an SEE of 4.5 dB. 4) Nevertheless, any PTA variation proposed in this study could be applied reliably to the prediction of SRT regardless of the audiometric patterns except for HFSS audiograms, as evidenced by correlation and regression analyses.

In conclusion, we confirmed that the most stable PTA variations that make the best-matched pairs with SRT for any audiogram are the conventional 3FA and W3FA doubling 1 kHz threshold. The addition of higher frequencies than 2 kHz to a PTA formula seems to impede the PTA-SRT agreement, especially for HFSS audiograms. In predicting SRT from PTA in a HFSS audiogram, W3FA should be the method of choice. AAO-HNS PTA is appropriate for atypical audiograms.

REFERENCES

1) Ventry IM. Pure tone-spondee threshold relationships in functional hearing loss: a hypothesis. J Speech Hear Disord 1976;41(1):16-22.
2) Harris JD, Haines HL, Myers CK. A new formula for using the audiogram to predict speech hearing loss. AMA Arch Otolaryngol 1956;63(2):158-76.
3) Graham JT. Evaluation of methods for predicting speech reception threshold. Arch Otolaryngol 1960;72:347-50.
4) Fletcher H. A method of calculating hearing loss for speech from an audiogram. Acta Otolaryngol Suppl 1950;90:26-37.
5) Coren S, Hakstian AR. Predicting speech recognition thresholds from pure tone hearing thresholds. Percept Mot Skills 1994;79(2):1003-8.
6) Delk JH, Glorig A, Quiggle RR, Summerfield AB. Predicting hearing loss for speech from pure tone audiograms. Laryngoscope 1957;67(1):1-15.
7) Carhart R. Observations on relations between thresholds for pure tones and for speech. J Speech Hear Disord 1971;36(4):476-83.
8) Chien CH, Tu TY, Chien SF, Li AC, Yang MJ, Shiao AS, et al. Relationship between Mandarin speech reception thresholds and pure-tone thresholds in the geriatric population. J Formos Med Assoc 2006;105(10):832-8.
9) Gurgel RK, Jackler RK, Dobie RA, Popelka GR. A new standardized format for reporting hearing outcome in clinical trials. Otolaryngol Head Neck Surg 2012;147(5):803-7.
10) Carhart R, Porter LS. Audiometric configuration and prediction of threshold for spondees. J Speech Hear Res 1971;14(3):486-95.
11) Ham TY. A study of Korean word lists for speech audiometry and Speech discrimination. Catholic Med J 1962;5:31-8.
12) Hahm TY. Complementary study on construction of Korean word lists for speech audiometry. J Inje Med Coll 1986;7(1):1-19.
13) Hederstierna C, Hultcrantz M, Collins A, Rosenhall U. Hearing in women at menopause. Prevalence of hearing loss, audiometric configuration and relation to hormone replacement therapy. Acta Otolaryngol 2007;127(2):149-55.
14) Demeester K, van Wieringen A, Hendrickx JJ, Topsakal V, Fransen E, van Laer L, et al. Audiometric shape and presbycusis. Int J Audiol 2009;48(4):222-32.
15) Kim J, Lee JH, Lee HJ, Kim HJ. Reliability of speech recognition threshold test using the Korean standard bisyllabic word list for adults in comparison with conventional Hahm's list. Korean J Otorhinolaryngol-Head Neck Surg 2012;55(6):350-4.
16) Kim HK, Lee CH. The correlation between speech reception threshold and pure tone audiometry. Korean J Otolaryngol-Head Neck Surg 1981;24(4):543-5.