J Am Acad Audiol 30:302–314 (2019)

Tracking of Noise Tolerance to Predict Hearing Aid Satisfaction in Loud Noisy Environments

DOI: 10.3766/jaaa.17101

Eric Seper*
Francis Kuk*
Petri Korhonen*
Christopher Slugocki*

Abstract

Background: A method that tracked tolerable noise level (TNL) over time while maintaining subjective speech intelligibility was reported previously. Although this method was reliable and efficacious as a research tool, its clinical efficacy and predictive ability of real-life hearing aid satisfaction were not measured.

Purpose: The study evaluated an adaptive method to estimate TNL using slope and variance of tracked noise level as criteria in a clinical setting. The relationship between TNL and subjective hearing aid satisfaction in noisy environments was also investigated.

Research Design: A single-blinded, repeated-measures design.

Study Sample: Seventeen experienced hearing aid wearers with bilateral mild-to-moderately-severe sensorineural hearing loss.

Data Collection and Analysis: Participants listened to 82-dB SPL continuous speech and tracked the background noise level that they could "put up with" while subjectively understanding >90% of the speech material. Two trials with each babble noise and continuous speech-shaped noise were measured in a single session. All four trials were completed aided using the participants' own hearing aids. The stimuli were presented in the sound field with speech from 0° and noise from the 180° azimuth. The instantaneous tolerable noise level was measured using a custom program and scored in two ways; the averaged TNL (aTNL) over the 2-min trial and the estimated TNL (eTNL) as soon as the listeners reached a stable noise estimate. Correlation between TNL and proportion of satisfied noisy environments was examined using the MarkeTrak questionnaire.

Results: All listeners completed the tracking of noise tolerance procedure within 2 min with good reliability. Sixty-five percent of the listeners yielded a stable noise estimate after 59.9 sec of actual test time. The eTNL for all trials was 78.6 dB SPL (standard deviation [SD] = 4.4 dB). The aTNL for all trials was 78.0 dB SPL (SD = 3.3 dB) after 120 sec. The aTNL was 79.2 dB SPL (SD = 5.4 dB) for babble noise and 77.0 dB SPL (SD = 5.9 dB) for speech-shaped noise. High within-session test–retest reliability was evident. The 95% confidence interval was 1.5 dB for babble noise and 2.8 dB for continuous speech-shaped noise. No significant correlation was measured between overall hearing aid satisfaction and the aTNL ($\rho = 0.20$ for both noises); however, a significant relationship between aTNL and proportion of satisfied noisy situations was evident ($\rho = 0.48$ for babble noise and $\rho = 0.55$ for speech-shaped noise).

Conclusion: The eTNL scoring method yielded similar results as the aTNL method although requiring only half the time for 65% of the listeners. This time efficiency, along with its reliability and the potential relationship between TNL and hearing aid satisfaction in noisy listening situations suggests that this procedure may be a good clinical tool to evaluate whether specific features on a hearing aid would improve noise tolerance and predict wearer satisfaction with the selected hearing aid in real-life loud noisy situations. A larger sample of hearing aid wearers is needed to further validate these potential uses.

Key Words: hearing aid satisfaction, reliability, tolerable noise level, tracking of noise tolerance

*Widex Office of Research in Clinical Amplification (ORCA-USA), Lisle, IL

Corresponding author: Francis Kuk, Widex Office of Research in Clinical Amplification (ORCA-USA), Lisle, IL 60532; Email: fkuk@widex.com

Eric Seper is now at the University of Chicago Medicine.

Abbreviations: ANL = acceptable noise level; aTNL = averaged tolerable noise level; CI = confidence interval; eTNL = estimated tolerable noise level; IOI-HA = International Outcome Inventory for Hearing Aids; LDL = loudness discomfort level; RMS = root mean square; SD = standard deviation; SNR = signal-to-noise ratio; TNL = tolerable noise level; TNT = tracking of noise tolerance

INTRODUCTION

Merely 40% of individuals with moderate to severe hearing losses and only 10% of those with mild hearing losses obtain hearing aids (Kochkin, 2010). Poor listening in noise has been cited as a factor limiting hearing aid uptake (Kochkin, 2000; Bertoli et al, 2009). Noisy backgrounds at moderate to loud levels can be uncomfortable in which to listen. In these environments, listeners may tolerate the discomfort only for a limited amount of time to understand what is said before they disengage themselves from the listening tasks. This disengagement could result in failed social interactions and perpetuate the isolation caused by a hearing loss (Hawthorne, 2008). If some hearing aids result in an even poorer tolerance in noisy backgrounds than the unaided condition (e.g., Kuk et al, 2017), it is reasonable to expect that they would not have been adopted; and when adopted, yield limited subjective satisfaction. Thus, clinicians may be interested in evaluating how well listeners tolerate background noise at the time of the fitting with the chosen hearing aids as a means to preemptively assess the wearers' potential real-life satisfaction of the hearing aids. In this study, we reported the results of a tool that measures a listener's aided tolerance for background noise while subjectively maintaining their understanding of continuous speech in a clinical setting. Furthermore, we were interested in evaluating the relationship between such a noise tolerance measure and subjective hearing aid satisfaction, particularly in loud, noisy environments.

Audiologists have used multiple methods to investigate listeners' responses to loud sounds. A classic form of loudness evaluation is the loudness discomfort level (LDL), which is the intensity level at which a listener considers the signal to be uncomfortably loud (Davis et al, 1946; Carhart, 1947; Watson and Tolan, 1949). This measurement is often completed unaided, using pure tones or a speech stimulus. Mueller and Bentler (2005) recommended establishing unaided LDL to ensure that hearing aid output does not exceed the listener's discomfort level. The unaided LDL is a popular clinical measurement as more than 80% of audiologists reported using LDL measurements at some point during their clinical practice (Mueller, 2003). Although frequency-specific LDL measurements have been recommended by diagnostic audiology and hearing aid fitting protocols (i.e., AAA, 2006), Beattie and Boyd (1986) obtained LDLs for pure tone and for speech stimuli and calculated the correlation between the two measures. The authors reported that pure tones were a poor predictor of loudness discomfort for speech stimuli because of weak (yet significant) correlation and high standard errors. Furthermore, Filion and Margolis (1992) reported that listeners may also judge signals used in the clinic (such as pure tones) differently than those encountered in the real world. Listeners with high unaided LDL can potentially tolerate more sound in noisy situations in theory; however, that measurement is completed without considering speech understanding. In real-life situations, satisfaction in noisy environments must consider both noise levels being tolerable and speech being intelligible. That is, the listener not only tolerates the noise, but also is functional in that environment. If the discomfort level for speech is desired, it is recommended that a specific measurement be made using speech materials, and not inferred from the pure tone data. In addition, pure tone LDLs do not account for loudness summation in a broadband signal, which means that the listener could still experience discomfort in certain loud environments. Finally, if a listener's LDL is low, a clinician could limit hearing aid maximum power output which could be detrimental toward speech recognition in noise at high inputs and be a source of hearing aid dissatisfaction (Kuk et al, 2011). Kochkin (2010) reported that "use in noisy situations" and "comfort with loud sounds" were among the areas that received the highest negative ratings in the MarkeTrak VIII hearing aid satisfaction survey. Because hearing aid use in noisy situations remains a concern for hearing aid wearers despite reported clinical use of the unaided LDL measurement, a clinical tool that specifically evaluates realistic loud stimuli, such as loud speech in the presence of background noise, may be of clinical value.

One method of evaluating a listener's response to loud sounds using speech stimuli instead of pure tones is to investigate loudness tolerance. The acceptable noise level (ANL) method proposed by Nabelek et al (1991) evaluates a listener's noise tolerance by measuring how much noise a listener is willing to "put up with" while speech is presented at the most comfortable listening level. The ANL measurement, compared with the LDL, may be a better predictor of hearing aid use, because the test evaluates a more typical real-world listening situation—speech stimuli in the presence of background noise. Although speech material in the ANL is presented at a comfortable level, listeners are not traditionally instructed to understand speech greater that a criterion level (i.e., 80%). The measurement is completed by first obtaining the most comfortable listening for speech, using a bracketing approach with 5 dB followed by 2-dB step sizes. Noise is then introduced and the clinician raises the level of the noise until the listener is no longer willing to "put up with" the noise level without becoming tired while listening. Nabelek et al (1991; 2006) measured ANL in the unaided condition and found that successful full-time hearing aid wearers tolerated background noise with an average signal-to-noise ratio (SNR) of 7.5 dB, which was more background noise tolerated than part-time hearing aid wearers (greater than 10 dB SNR). Nabelek et al (2006) later reported no difference between unaided and aided ANL performance. The authors also reported that listeners with ANL lower than 7 dB SNR were likely successful hearing aid wearers; listeners with ANL greater than 13 dB SNR were likely unsuccessful hearing aid wearers. Furthermore, the study concluded that a listener's noise tolerance could predict his/her overall hearing aid success (as opposed to success in specific environments) with 85% accuracy. In this case, hearing aid success was defined as hearing aid usage (frequency of hearing aid use).

The ANL has also been used to demonstrate the efficacy of hearing aid processing. Specifically, hearing aid processing can improve aided ANL (greater tolerance for background noise) and increase potential hearing aid success. If hearing aid processing can improve ANL to 7 dB or better [criteria of successful hearing aid use set forth by Nabelek et al (2006)], then one can predict that the listener will be a successful hearing aid wearer. For example, Freyaldenhoven et al (2005) reported a 3.5 dB ANL improvement from the use of a directional microphone when speech was presented from the front and noise from the back. Peeters et al (2009) reported that a noise reduction algorithm and the directional microphone improved ANL by 3.3 dB and 2.8 dB, respectively. Wu and Stangl (2013) reported that wide dynamic range compression hearing aids worsened ANL by 1.5 dB (compared with linear) whereas noise reduction and directional microphone improved ANL by 1.1 dB and 2.8 dB, respectively. Olsen and Brännström (2014) provided a summary of the studies conducted on the topic of ANL between 1991 and 2012.

Studies have also examined the relationship between noise tolerance and subjective hearing aid satisfaction or benefit. Taylor (2008) reported a significant relationship between unaided ANL and the International Outcome Inventory for Hearing Aids (IOI-HA, Cox and Alexander, 2002) questionnaire. Later, Ho et al (2013) reported significant correlation between the unaided ANL and the IOI-HA questionnaire using Taiwanese speech materials; however, the authors cautioned that the prediction of hearing aid success was lower than that reported by Nabelek et al (2006). On the other hand, Freyaldenhoven et al (2008) reported that both unaided and aided ANLs were not related to any subscales on the Abbreviated Profile of Hearing Aid Benefit, including Ease of Communication, Background Noise, and Aversiveness to Sounds (Cox and Alexander, 1995). Olsen et al (2012) could not report any association between ANL and subjective hearing aid outcome using the IOI-HA questionnaire. Previous research has been inconclusive toward determining the relationship between noise tolerance and subjective hearing aid satisfaction. Therefore, it is worthwhile to continue exploring this relationship, but with different strategies (i.e., noise tolerance test materials, procedures, and questionnaires) to determine if a separate measure could demonstrate a relationship between noise tolerance and subjective hearing aid satisfaction.

Olsen et al (2012) supported the Nabelek et al (1991; 2006) findings and reported a better ANL for full-time hearing aid wearers (in contrast to part-time wearers). Additional studies replicated some, but not all, elements of the Nabelek et al (1991; 2006) studies. For example, Walravens et al (2014) reported similar ANL values for full-time hearing aid wearers as Nabelek et al (1991; 2006). However, in contrast to the Nabelek et al (1991; 2006) studies, part-time hearing aid wearers or non-hearing aid wearers produced better ANL values compared with the full-time wearers. Wu et al (2016) did not replicate the association between ANL and hearing aid usage from Nabelek et al (2006). A possible explanation for the variability in outcomes among studies could be the repeatability of the ANL procedure. Olsen et al (2012; 2013), while using Danish passages as stimuli, showed intra-tester variability to be 6.5–8.6 dB for sessions completed within the same day and 7.1–8.8 dB for sessions completed on separate days in hearing-impaired participants. The authors concluded that variability could be reduced to 4 dB after running three trials of the ANL. However, the test's repeatability is a potential limiting factor.

What do ANL and LDL truly measure? Franklin et al (2012) measured ANL and LDL for normal hearing individuals in the sound field. They reported similar LDL and ANL values to those reported in prior works, but also reported a lack of correlation between ANL and the LDL measurement and suggested that loudness discomfort (i.e., LDL) and loudness tolerance measures (i.e., ANL) capture different attitudes about noise. As suggested in previous reports (i.e., Mueller and Bentler, 2005; Franklin et al, 2012), the LDL may be clinically useful in determining hearing aid output limits. Because subjective satisfaction with hearing aids in loud, noisy environments is likely influenced by both comfort and intelligibility, the LDL may not be the best tool to evaluate hearing aid satisfaction. A noise tolerance measure, on the other hand, may offer insight toward the hearing aid wearer's potential satisfaction with a hearing aid and the hearing aid's various processing. Considering the potential test–retest reliability concerns [i.e., Brännström et al (2014)] with the ANL and disagreement regarding the correlation between ANL and subjective hearing aid satisfaction, another method that reports high test–retest reliability is desirable.

Kuk et al (2017) reported on a method that tracked noise tolerance over time in a research setting. The tracking of noise tolerance (TNT) test tracked the listeners' noise tolerance while requesting them to maintain a criterion level of subjective speech understanding. The research TNT presented speech (discourse passage) at 85 dB SPL in the presence of a continuous speech-shaped noise over a 4-min period. Listeners were instructed to increase the noise level using a touch screen apparatus until they could no longer "put up with" the noise or they noticed a decrease in their ability to understand the speech <90%. The TNT program tracked instantaneous noise tolerance over time, which could be beneficial in ensuring a stable estimate of noise tolerance because momentary fluctuations in loudness judgments were averaged. Results obtained by Kuk et al (2017) showed that the TNT procedure had a within-session test–retest variability of 2.2 dB (95% confidence interval [CI]) averaged across hearing aid conditions for thirteen experienced hearing aid wearers. The TNT also offered a means to compare hearing aid algorithms where differences in output levels were expected, especially where these algorithms differed in their time course of action. Kuk et al (2017) demonstrated that digital noise reduction improved noise tolerance by ~3 dB. Listeners reached a stable noise tolerance level sooner when using a hearing aid with a faster noise reduction adaptation rate than one with a slower adaptation time even though both hearing aids provided the same amount of gain reduction in noise. The authors also reported ~4.5 dB improvement in noise tolerance with a directional microphone compared with an omnidirectional microphone. When both the directional microphone and noise reduction algorithms were activated, aided hearing-impaired listeners tolerated similar noise levels compared with normal hearing listeners.

In the current study, we were interested in the TNT's ability to predict hearing aid satisfaction in real-world loud, noisy environments. A possible link between the two may exist because it is reasonable to expect that individuals who can tolerate a high level of noise are more likely to stay longer in the noisy environments, and thus be more satisfied with the performance of the hearing aids in that environment than ones that do not allow such tolerance. To accomplish that task, we believed the TNT should be investigated in a clinical setting, where the merits of the TNT reliability and sensitivity can be demonstrated, even if available clinical time is limited. We believed that a few changes to the original TNT test might demonstrate the *clinical* acceptability of the test. The first change promoted active test participation through a Bekesy-style automatic noise tracking paradigm (where the noise level automatically changes at a specified rate) with the intention that it may lead to a faster accurate stable noise level estimation. In the Kuk et al (2017) study, the noise level began at 75 dB SPL (+10 dB SNR) and stayed at the same level until the listener manually increased (or decreased) the level until they could "just put up with" the noise while still understanding speech. Stephens and Anderson (1971) investigated the differences between a manual method of level adjustment, such as the approach used in Kuk et al (2017), and a Bekesy-style tracking system when measuring LDL. The authors determined that when using a manual approach, naïve listeners had 10 dB lower LDL values compared with the Bekesy tracking procedure. Morgan et al (1974) further reported ~7 dB lower LDL using the manual method compared with the Bekesy tracking method. It is possible that a Bekesy method of adjustment delivers more stable results, because the rate of noise level change is predetermined and not dependent on the clinician's or the listener's response rate. The lack of consistency for the rate of level change would be controlled using the Bekesy procedure. This systematic procedure could deliver stable results more quickly than the manual method. Morgan et al (1974) also reported that LDL values increased over the span of six trials with the manual method, whereas Bekesy tracking resulted in stable results across all six trials. Thus, we modified our previous TNT measure so that the noise level automatically changed at a specific rate set by the clinician, with the listener continuously controlling the noise level changes using a computer keyboard.

A second update to the TNT was to shorten the time required to complete the test. We recognized that some listeners were able to reach a stable noise tolerance level within a minute; whereas others may require as much as 2 min to reach a stable level. Thus, we introduced an adaptive scoring method which estimated the stability of the noise tolerance level (and the time required to reach that stable level). The tracking would terminate once the criteria for a stable estimate were met. This procedural change was beneficial for two reasons. First, it provided an objective, systematic, and consistent means of estimating the stable noise level free from any inadvertent clinician biases (i.e., from different clinicians at different sites or the same clinician at different times), especially if their tolerable noise levels (TNLs) varied greatly over the course of the test. Second, because the program could terminate immediately (instead of continuing for the whole 2 min) once the listener reached a stable estimate, valuable clinical evaluation time could be saved for some listeners.

Finally, it was desirable to include speech materials specifically dedicated to the TNT. These dedicated materials included a consistent topic throughout the passage to replicate a realistic communication situation.

Speech intelligibility is a criterion for the listener when completing the TNT and any abrupt changes on the topic could influence speech intelligibility judgment. That is, changes to the topic of conversation without warning are rare for real-world communication; thus, we wanted to avoid this unlikely real-world scenario in the testing. Simple, common-knowledge topic speech materials with enough length in each passage to span an entire TNL trial likely provide a reliable estimate of the TNL because topic continuity is maintained throughout the test.

The intent of the current study was to evaluate the relationship between the TNL and subjective hearing aid outcomes in a clinical population. For that reason, one objective was to characterize the efficiency of the online estimation procedure and its reliability in a clinical population. A second objective was to examine if the stable noise level correlated with any subjective impression of hearing aid performance.

METHODS

Participants

Data were collected at two clinical sites from 17 experienced hearing aid wearers. These sites were private audiology clinics serving adult patients in Canada (see Acknowledgments). Consented, regular hearing aid patients who returned for routine hearing aid follow-up services were recruited. There were nine participants from site 1 and eight participants from site 2. The four-frequency pure tone average (500, 1000, 2000, and 4000 Hz) for the first site was 49.2 dB HL (across both ears) and 43.4 dB HL for the second site. Figure 1 shows the combined mean audiometric thresholds of the participants from the two sites. The average age of the participants was 69.7 years (standard deviation [SD] = 10.5 years) for the first site and 75.6 years (SD = 3.3 years) for the second site. Participants were not compensated for their participation.

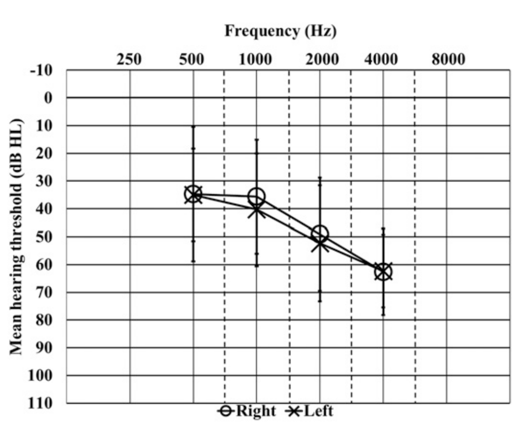

Figure 1. Mean hearing thresholds at 500, 1000, 2000, and 4000 Hz for the 17 hearing-impaired participants. Error bars indicate ±1 SD.

Hearing Aids

Participants used their own hearing aids during the TNT testing. They included hearing aids from three manufacturers—Phonak (5 wearers), Siemens (1 wearer), and Widex (11 wearers). Fifteen of the 17 participants wore a receiver-in-the-canal style hearing aid, one participant used behind-the-ear hearing aids with custom ear molds, and one participant used custom in-the-ear style hearing aids. All the hearing aids were digital devices with three classified as "premium," seven as "mid-level," and seven as "entry-level" digital technology by their manufacturers. All the hearing aids used multichannel wide dynamic range compression fitted based on the National Acoustics Laboratory Non-linear 2 (NAL-NL2) targets (Keidser et al, 2011). All included directional microphones and noise reduction algorithms. Participants also completed the MarkeTrak questionnaire (Kochkin, 2010) to reflect their degree of satisfaction for their own hearing aids. This questionnaire was chosen for its simple and quick administration and its coverage over many aspects of hearing aid use (such as satisfaction with sound quality, use in multiple listening environments, ergonomic factors, and quality of life assessment).

Development of the Revised Clinical TNT

Preparation of Stimuli

Speech Stimuli: The TNT test included seven passages each on a common topic (e.g., movies, birds, music) and spanning 2 min in length. The material was adapted from the simple English version of Wikipedia. This version of Wikipedia is written in simple English using rudimentary vocabulary and grammar. The text from Wikipedia was altered to remove long sentences and clauses. The passages were written to reflect as near a fifth grade reading level as possible based on the Flesch–Kincaid Grade Level score (Kincaid et al, 1975). Because speech intelligibility is a criterion for the listeners' determination of TNL, and a particularly difficult section of speech material could reduce the listener's noise tolerance via decreased intelligibility, it is important to maintain a similar level of reading difficulty throughout the speech passage. The reading level grade score was compared between the first half of the passage (2 min and 15 sec) and the second half of each passage, with a goal of remaining near a fifth grade reading level.

Speech stimuli were recorded in an audiometric sound booth (3 × 3 × 2 m). A male talker with a Midwestern American English dialect sat in the middle of the booth. A large diaphragm (1 inch) condenser microphone placed ~1–2″ directly in front of the talker was used for all recordings. Signal from the microphone was routed via a Mackie mixer to Echo Firewire12 audio interface. The audio software used in capturing the signal was Adobe Audition 1.0. All the materials were recorded using 44.1 kHz sampling frequency and 32-bit accuracy.

Each speech passage was edited to have a similar intensity level (average root mean square [RMS]) throughout the duration of the passage. The talker spoke at a normal vocal effort and monitored his production throughout the recordings using a sound level meter. The average RMS level of the recordings was equalized subsequently in 1-min segments. The passages were then edited for a "shouted" version using the SII Standards (ANSI S3.5, 1997). We applied the difference in spectra between the two vocal efforts (normal versus. shouted) described in the Standards to the normal speech file at one third-octave bands using a custom filter made in Adobe Audition. The resulting stimuli reflected greater high frequency emphasis in the shouted version relative to the normal version at approximately 65 dB SPL.

Noise Stimuli: The test included two types of noise: continuous speech-shaped noise and a 4-talker babble noise with two streams. The continuous speech-shaped noise was generated for each individual speech passage to have the same spectral shape as the long-term average spectrum of each speech passage. We determined the 80th order linear predictive coefficients of the speech signals and applied those coefficients to design a filter. The input to this filter was white noise. After filtering, the output noise contained the same spectral envelope as the long-term average spectrum of the speech materials.

The source material for the babble noise was from two male and two female talkers. The passages used in the generation were obtained from audiobooks in the public domain. The original recordings were carried out using 44.1 kHz audio sampling frequency. Each passage was 30 sec long and was equalized for max RMS level in a 50 ms sliding window. The final babble noise consisted of eight streams (4 talkers $\times$ 2 streams [i.e., same talker but time delayed by 5 sec]).

TNT Test Flow

The program presented a constant-level 82 dB SPL speech, which is loud speech that would be typically encountered in a loud, noisy background (Pearsons et al, 1977). Ongoing background noise automatically increased in level (i.e., 0.1 dB per 0.1 sec) until the listener pressed the spacebar on a keyboard to decrease the noise level. The noise level continued to decrease as long as the listener held the spacebar down. Once the spacebar was released, the noise level increased. The rate of decrease in level was the same as the increase in level. Each tracking trial (speech and noise) lasted for 2 min. The program graphically displayed the TNL over time on the clinician's computer monitor. The average level of the TNL tracked over the 2-min period was reported as the aTNL.

TNT Programing and Algorithms

The TNT test software was implemented for the Microsoft Windows operating system (Compatible with Windows XP up to at least Windows 10). The TNT program was developed using VB.net programing language. In addition to reporting the aTNL, we added an adaptive algorithm that tracked the TNL and reported on the time once a stable TNL was estimated. This was performed to shorten the test time (i.e., increase efficiency) for listeners who may reach a stable tolerance level in $<$2 min. This algorithm tracked the slope (change in noise level over time) and variance (how much the noise level deviated from the mean noise level) of the changing noise every 0.1 sec. Ideally, the slope and variance of the noise tracking function at the stable state would be zero. In practice, the variance would not be zero for any period greater than 0.1-sec segments, because the noise level automatically changes every 0.1 sec. We set a criterion such that the program could terminate when the slope was 0 ± 0.05 and the variance $<$2 dB for a 30 s period. The noise level estimated using this algorithm was called the estimated TNL (eTNL). The program displayed this estimate and the time at which that stable level was reached. Although the program was capable of terminating after eTNL was determined, all participants completed the entire 2-min test. Thus, the program displayed both the aTNL that was calculated across the entire 2-min test and the eTNL calculated as soon as a stable noise estimate was made. All results were saved in a text file used in data analysis.

Equipment and Setup: Both of the participating audiology clinics installed the TNT software on their existing clinical computers. Sites used existing clinical equipment, including a clinical audiometer, sound booth, and sound field loudspeakers (Edirol MA-7A loudspeakers at one site; Grason-Stadler, Inc. [GSI] loudspeakers at the other site) placed at 0° and 180° azimuth relative to a chair where the participants were seated. The speech and noise stimuli were routed from the computer through the GSI-61 clinical audiometer using stereo audio cables and the "external audio input" jack to the audiometer. A calibration stimulus (speech-shaped noise used in the TNT program) was included in the TNT software. The sites were instructed to set the attenuators on the audiometer for both channels to 67 dB HL for a desired output level of 82 dB SPL. Next, the calibration stimulus was presented and the VU meter set to peak at zero. Both clinics used the SoundTracker feature in Widex Compass GPS computer software to calibrate the stimulus levels. This method yielded similar measurement (within 1–2 dB) as a sound level meter (Kuk et al, 2004) and real ear measurement systems (85% within 2 dB; Oeding and Valente, 2013).

Procedure: Before TNT testing, all participants completed the MarkeTrak questionnaire (Kochkin, 2010) to report their satisfaction with their current hearing aids in different listening situations. They rated their satisfaction using a 7-point Likert scale, where "1" was "very dissatisfied," "4" was "neutral," and "7" was "very satisfied." All participants completed the TNT with 82 dB SPL speech presented from 0° azimuth and noise from 180°. This loudspeaker arrangement allowed the possibility for evaluation of hearing aid features that required spatial separation between speech and noise sources [e.g., directional microphones, Freyaldenhoven et al (2005)]. The starting level of the noise was set at a SNR = +10 dB. A practice test using babble noise was administered before beginning the actual trials. Each participant completed the TNT with babble noise and continuous speech-shaped noise in a counterbalanced order. Each clinic tested both noise conditions twice. Speech lists and noise types were counterbalanced to reduce any potential order or list effects. The following instructions were verbally provided to each participant before the practice TNT trial:

Acceptability of Speech Level

> "You will be listening to a male talker reading a story at a loud volume. You should find this volume to be loud, but not uncomfortably loud. You should also be able to understand ≥90% of the words. If this is not the case—either not understanding enough or too loud—please let me know and I will adjust it accordingly."

At this point, if the stimulus was too loud, the test administrator was instructed to lower the level to a "loud, but acceptable level." If the listener could not understand ≥90% of words, the administrator was instructed to see if reducing or increasing the speech level would help. If not, the test administrator was instructed to note the highest level of understanding reached and stay at that speech level. None of the participants included in the study required changes in speech level.

Determining TNL

> "You will hear some noise in the background while you listen to the male talker. The noise will automatically get louder. I want you to monitor the noise level and maintain the loudest noise level you can put up with while still understanding 90% of the words in the story. If the noise becomes too loud, where you can no longer put up with it or understand <90% of the words in the story, you can turn the noise down by pressing and holding the space bar. If it appears softer than before, you should allow the volume to increase by letting go of the space bar. If it is louder than before, you should turn the volume down to keep at the same level by pressing the space bar again. Your ability to understand speech should never change to <90%. The test will run for 2 min and then stop."

It should be noted that the same TNT instructions were used in the Kuk et al's (2017) study. It was demonstrated then that listeners' objective speech recognition scores at the TNL exceeded the 90% subjective criterion that was set in the instructions. That supported the speculation that listeners included a judgment of speech intelligibility as one of their criteria in tracking tolerable noise. In addition, they were able to meet that criterion during the tracking. Thus, no further validation of speech intelligibility was conducted in this study.

RESULTS

Evaluating the TNT as a Clinical Procedure

Comparison between TNL Scoring Methods

Two methods of estimating the TNL were used in the study. One method averaged the TNL over the entire 2-min test (aTNL). The second method eTNL from the slope and variance calculation for a stable 30 sec noise tracking. The motive behind the second method was to determine if additional test time may be saved from listeners who were stable in their responses. We were interested in determining the difference between scoring methods, independent of noise type. Thus, the results for both babble noise and continuous speech-shaped noise were combined when evaluating the difference between scoring methods. Although all 17 participants completed all TNL trials with both noise types, not all participants were able to reach a stable eTNL within the 2-min test window based on the variance criterion that we set (<2 dB). In all, the eTNL could not be established for five participants when babble noise was used and for six participants when continuous speech-shaped noise was used. Of the six participants who did not establish eTNL for continuous speech-shaped noise, five of them also did not establish eTNL for babble noise. The data from all 17 participants were included in the aTNL calculation. Of the listeners who stabilized within the 2-min tracking period, a stable eTNL was estimated at 78.6 dB SPL (SD = 4.4 dB) after an average of 59.9 sec (SD = 18.9 sec) of actual test time. The aTNL was 78.0 dB SPL (SD = 3.3 dB). The difference between the eTNL and the aTNL, when considering only those participants who achieved a stable eTNL, was not significant as reflected on a paired sample, two-tailed t-test [$t_{(16)} = 1.33$, $p = 0.20$].

Effect of Noise Type

As reported earlier, the eTNL was not established for up to six participants because of variance that exceeded the criterion for stable estimate. The following results compared only the aTNL data because they included all the participants. Figure 2 reports the average noise levels for the babble noise trials and the average noise levels for the continuous speech-shaped noise trials over the 2-min TNT test. The aTNL was 79.2 dB SPL (SD = 5.4 dB) for babble noise and 77.0 dB SPL (SD = 5.9 dB) for continuous speech-shaped noise. The effect of noise type was significant [paired-sample, two-tailed t-test, $t_{(16)} = 4.49, p < 0.001$]. In other words, participants tolerated 2.2 dB less continuous speech-shaped noise than babble noise without a change in subjective speech understanding.

Test–Retest Reliability

We calculated the within-session test–retest difference in aTNL for each individual hearing-impaired listener across both noise types. Within-session variability was estimated by determining the difference in TNL between trial 1 and trial 2 for each noise type. The mean and standard error of the difference in aTNL between trials were then used to calculate the 95% CI for the difference (Table 1). Test–retest reliability of aTNL, as measured with Pearson's correlation, was high for both babble ($r = 0.91$, $p < 0.001$) and continuous speech-shaped ($r = 0.78$, $p < 0.001$) noise (Figure 3). The 95% CI was also observed to be smaller for babble (1.5 dB) relative to continuous speech-shaped noise (2.8 dB).

TNL and Hearing Aid Satisfaction

Figure 4 shows the overall satisfaction ratings on the MarkeTrak questionnaire against the aTNL for all participants. All participants reported that they were satisfied (rating of 5 or higher) when asked, "Overall, how satisfied are you with your hearing aids?" Nearly all (16 of 17) of these satisfied hearing aid wearers achieved aTNL greater than 72 dB SPL for babble noise and 67 dB SPL for continuous speech-shaped noise. One participant tolerated 86.9 dB SPL of continuous speech-shaped noise, whereas a different participant with the same satisfaction rating tolerated only 68.7 dB SPL of continuous speech-shaped noise. These results revealed a difference of up to 18.2 dB of TNL at the "7—very satisfied" rating. No significant correlation existed between overall satisfaction rating and TNL level ($p > 0.05$).

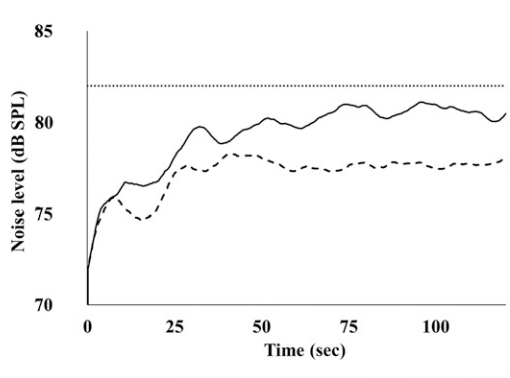

Figure 2. Average TNLs for hearing-impaired participants listening in babble noise (solid line) and continuous speech-shaped noise (dashed line). Speech level (82 dB SPL) indicated by dotted line.

To further investigate if TNL values would correlate with listening in loud, noisy situations, we identified nine specific listening situations on the MarkeTrak questionnaire that concerned listening in challenging, often noisy, situations and examined their ratings to the aTNL. The nine situations were: "while following a conversation in noise," "listening in large groups," "listening while shopping," "listening at a large lecture hall," "listening at a place of worship," "listening in a restaurant," "listening in a car," "listening to loud voices," and "listening on a noisy street." Because not all participants encountered all nine listening situations, we calculated the proportion of satisfied loud, noisy listening conditions for each participant as a ratio between the total numbers of satisfied loud noisy conditions to the number of loud noisy conditions experienced by the individuals. Of the 17 participants, seven encountered all nine situations, three encountered eight situations, two encountered seven situations, three encountered six situations, and two encountered five situations. All the participants encountered "while following a conversation in noise." There were typically one or two participants who did not encounter a specific noisy situation. The exceptions were the items "listening at a large lecture hall" and "listening at a place of worship" where seven and five listeners had not encountered the situations, respectively. In our scoring, a rating of "5 or relatively satisfied" or higher was considered a "satisfied" response, whereas a rating of "4" or lower was considered a "dissatisfied" response. Figure 5 reports the measured aTNL plotted against the proportion of loud, noisy listening conditions in which participants reported satisfaction with their hearing aids. An outlier analysis was performed before any further statistical analysis. The result showed that all but one listener's performance were within a ±2 SD criterion. Thus, all the data were included in the subsequent analyses. In addition, because the data were not distributed normally, a nonparametric Spearman correlation was used to examine the relationship between TNL and proportion of satisfied rating. The results suggested that a higher aTNL value was significantly correlated with a greater proportion of loud, noisy listening conditions that were deemed satisfactory

Table 1. Mean, Standard Error (SE), Range, and 95% CIs for the aTNL Method Measured for Aided Hearing Impaired Participants

	aTNL Mean Test–Retest Difference	aTNL SE	Range	Lower	Upper	95% CI
Babble	1.6	0.4	(0.0, 5.1)	0.9	2.4	1.5
Continuous	2.9	0.7	(0.0, 8.7)	1.5	4.3	2.8

(babble noise, Spearman's $\rho = 0.48$, $p < 0.05$; continuous speech-shaped noise, Spearman's $\rho = 0.55$, $p < 0.05$).

When one examined the distribution of data on Figure 5, one cannot but notice that eight listeners were 100% satisfied in all their applicable, loud, noisy situations on the MarkeTrak questionnaire. This created a ceiling effect and could have lowered the calculated correlation between TNL and proportion of satisfied listening situations. We explored that possibility by correlating only the data on listeners who had a proportion of satisfied listening situations that was <1 (i.e., removing the data on listeners who reported total satisfaction). The results showed a stronger correlation (babble noise, Spearman's $\rho = 0.87$, $p < 0.01$; continuous speech-shaped noise, Spearman's $\rho = 0.77$, $p < 0.01$) between TNL and satisfaction in loud noisy situation for those who were not completely satisfied with their hearing aids in all the listed noisy situations. A caveat is that this analysis was based on an even smaller sample size (n = 9) and should only be viewed as exploratory. A larger sample of participants that includes various levels of hearing aid satisfaction would be needed.

DISCUSSION

The current study showed that the clinical version of the TNT test was reliable and may be sensitive to predict real-life satisfaction with the hearing aids in loud, noisy communication situations. Specifically, it was determined that both methods (estimation and averaging) of scoring TNL yielded similar values. This suggests that the estimation method can reliably estimate the listener's TNL in a shorter period of time (1 min or 50% of real-time savings) for listeners who are reliable and stable in their loudness judgment. The reliability of the clinical TNT method was similar to that reported in Kuk et al (2017). Most importantly, the current results revealed a potential relation between aTNL and subjective hearing aid satisfaction in loud noisy environments.

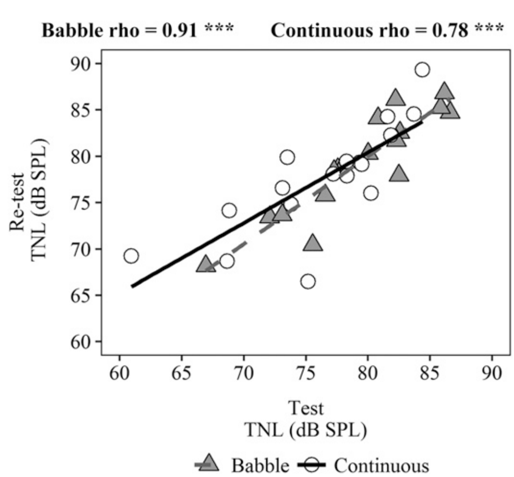

Figure 3. Test–retest reliability for 17 participants using babble and continuous speech-shaped noises for the aTNL thresholds (dB SPL).

How Would the TNL be Scored?

The current study showed that the reliability of the TNT procedure in estimating the TNL was high. Its efficiency can be further improved in listeners who reached a stable estimate sooner. Indeed, eleven (of 17) participants completed the TNL for both noise types within 1 min with an average nonsignificant difference between the eTNL and aTNL of 0.6 dB (eTNL greater than aTNL). This suggests that the two estimation methods yielded a similar TNL value. For maximum efficiency and reliability, the clinical TNT procedure would be conducted in the same manner for all listeners with a default tracking duration of 2 min. During the tracking, an online estimation of the slope and variance of the listener's tracking is performed. Tracking terminates as soon as the slope and variance criteria for stable performance are met. Listeners who are reliable in their tolerance estimate will likely meet the criteria within the 2 min and save evaluation time. Listeners who do not satisfy the criteria for the eTNL would complete the entire test duration of 2 min. This adaptive procedure could reduce the testing time for nearly 65% of the participants. The saved time becomes crucial if clinicians are interested in completing multiple comparisons to evaluate the effect of hearing aid processing, such as noise reduction or directional microphone. The saved time could also result in less effort (and possibly less fatigue) from the listener.

Which Noise Type to Use?

The TNT program allows the clinicians to evaluate noise tolerance for babble noise and continuous speech-shaped noise. The current study revealed a difference in noise tolerance between the two noise types. The listeners tolerated less continuous speech-shaped noise than

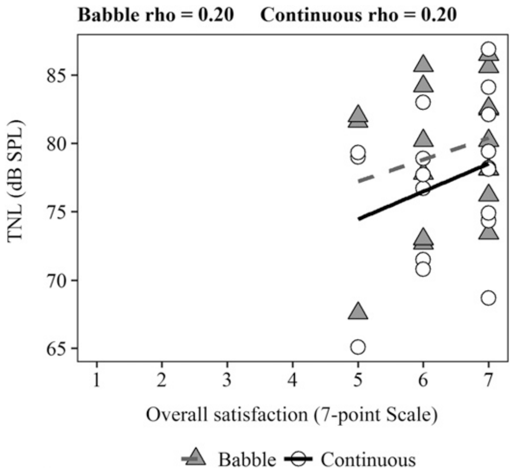

Figure 4. Overall Satisfaction rating on MarkeTrak questionnaire (7-point scale) plotted against TNL (dB SPL) for babble and continuous speech-shaped noise types.

babble noise by 2.2 dB. The results of the current study support Freyaldenhoven et al (2006), who also reported approximately 2 dB of noise tolerance difference between babble noise and continuous speech-shaped noise during the ANL test. With a speech-shaped noise that shares the same long-term spectral characteristics as the speech signal (such as the continuous speech-shaped noise used in the current study), energetic masking can result in a rapid decrease in intelligibility as SNR decreases (Brungart, 2001). On the other hand, Brungart (2001) reported that the effects of this energetic masking occur more gradually when the noise is spectrally different from the speech (such as the babble noise used in the current study), potentially because of the listener hearing a glimpse of the speech signal during the "valleys" of the modulated noise. Cooke (2006) reported that this "glimpsing" between the modulations provides enough information to support speech recognition. Therefore, a lower TNL is expected when using speech-shaped noise compared with babble noise, because poorer speech intelligibility with speech-shaped noise could drive the noise levels lower.

The different noise types may be useful for different purposes. Previous studies have selected speech babble as their noise stimulus (i.e., Freyaldenhoven et al, 2006; Nabelek et al, 2006) to be more representative of daily listening. The current study showed that babble noise was less variable (CI = 1.5 dB) in tracking the TNL compared with continuous speech-shaped noise (CI = 2.8 dB). Considering its representativeness of daily listening and the reliability measures, babble noise is an appropriate noise type for evaluating a listener's noise tolerance using the TNT procedure.

On the other hand, continuous speech-shaped noise may be a more useful noise type than babble noise when evaluating specific hearing aid features such as noise reduction. This is because most noise reduction algorithms recognize a continuous signal, and not a modulated signal, as noise. If the signal is modulated, the noise reduction algorithm may consider that as a "speech" signal and not reduce gain for this input (Bentler et al, 2008; Krishnamurti and Anderson, 2008). Comparisons between different types of noise reduction systems may be more efficient and demonstrative using an unmodulated noise type. Thus, we recommend continuous speech-shaped noise unless there are reasons to consider otherwise (such as claims that the noise reduction algorithm recognizes modulated signals as noise).

Test–Retest Reliability

The within-session test–retest reliability, as estimated with the 95% CI, was 1.5 dB for babble noise and 2.8 dB for continuous speech-shaped noise when the noise levels were averaged across the complete 2 min test. This is similar to what was reported in Kuk et al's (2017) original TNT test of 2.2 dB. A test that is highly correlated between trials is especially important when two or more measurements are necessary to demonstrate a difference in feature benefits. For instance, the first trial evaluates the hearing aid without a feature (i.e., noise reduction off) and the second trial evaluates the hearing aid with the feature turned on. If the difference in TNT between the noise reduction "On" and "Off" is greater than the magnitude of the within-session test–retest reliability, one may interpret that result as significant. Based on reliability measurements in the current study and results of Kuk et al (2017), the clinical TNT should be capable of allowing one to evaluate hearing aid feature benefit that exceeds 2 dB.

Predicting Hearing Aid Satisfaction

The results of the current study show that TNT may predict how hearing aid wearers respond to loud noisy

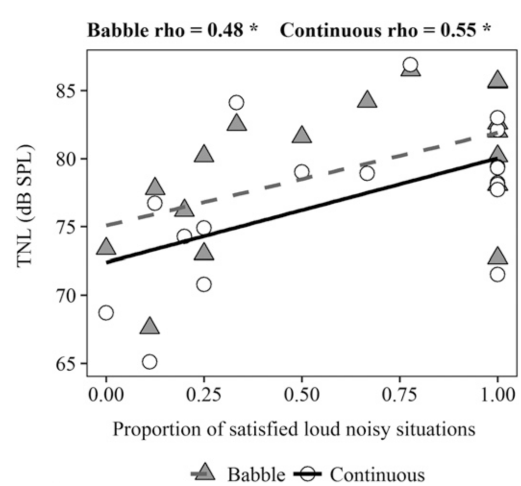

Figure 5. TNL (dB SPL) against proportion of loud, noisy situations that were reported as satisfied for babble and continuous speech-shaped noise.

environments. Listeners who can tolerate greater amounts of noise are more likely to be willing to spend more time and be satisfied in more loud noisy environments. Listeners who tolerate only a low level of background noise may be reluctant to spend time (and thus be dissatisfied) with the hearing aids in loud noisy situations. Increased engagement in social functions is an important aspect of auditory rehabilitation that audiologists promote. The current study showed that a higher TNL corresponded with satisfaction in a greater proportion of noisy listening situations. Those listeners who tolerate relatively low background noise levels may require hearing aid programing adjustments, such as lowering gain for loud inputs, lowering maximum power output, or selecting advanced features such as noise reduction or directional microphones to promote full-time and successful hearing aid usage (Dillon, 2001; Jenstad et al, 2003; Kuk et al, 2017). Kuk et al (2017) reported that use of noise reduction and directional microphone resulted in similar TNL values between hearing-impaired and normal-hearing listeners. Audiologists can also provide useful counseling for listeners with low TNLs such as recommending that listeners move conversations to quieter areas whenever possible. This could potentially decrease the amount of loud noise and improve speech intelligibility in a social setting. Furthermore, if listeners are novice hearing aid wearers, it may be important to remind them that normal hearing individuals also experience difficulty understanding speech in loud noise, and that environmental noise may become more acceptable with continued hearing aid usage (Dillon, 2001; Philibert et al, 2005; Dawes et al, 2014; Kuk et al, 2015). The current study did not find a relationship between TNL and overall hearing aid satisfaction. One reason is that all participants were satisfied wearers of their own hearing aids. This ceiling effect could have confounded the observation. In addition, the current study used a simple two loudspeaker setup (front and back) whereas real-world noise would be more diffuse. A more diffuse noise setup may predict a higher correlation with real-world satisfaction; however, such a setup may not be as realistic in the typical clinic. Lastly, overall satisfaction with hearing aids is likely the culmination of several subjective factors. Physical comfort, ease of use, listening in quiet situations, and device cost are among the factors that could contribute to a hearing aid wearer's attitudes regarding his/her hearing aids. Satisfaction of the hearing aids in loud noisy situations is an important determinant of overall hearing aid satisfaction, but not the only factor that predicts overall hearing aid success as reflected in the data shown in this study.

A limitation of the current study is that only a small sample size was used (n = 17). This resulted in a limited dispersion of data as only hearing-impaired participants who were satisfied with their hearing aids were enrolled in the study. This was not intentional as participating sites were instructed to recruit participants who were interested in the study regardless of their hearing aid satisfaction. Although the current study showed a moderate correlation between the TNL and the number of noisy environments in which the participants were satisfied, a correlation between TNL and overall MarkeTrak satisfaction was not found, potentially because of the absence of dissatisfied wearers, among other factors. Another area where the limited diversity of data may have affected our analysis was seen in the correlation between TNL and proportion of satisfied situations (i.e., Figure 5). Eight of seventeen listeners were satisfied in all their relevant listening situations (i.e., proportion of 1) whereas fewer listeners were represented with different degrees (or proportion) of satisfaction. This ceiling effect could have lowered the "true" correlation. This was a real possibility when we removed the data of listeners with 100% satisfaction (i.e., proportion of 1) in the correlation analysis. The resulting correlation coefficients improved from $\sim$0.5 to $>$0.8. A future study could explore the relationship between noise tolerance and hearing aid satisfaction with a greater number of participants at each level of hearing aid satisfaction. This will result in a more definitive conclusion on the relationship between TNL and hearing aid satisfaction in loud, noisy situations.

CONCLUSION

The current study reported that the TNT procedure is a reliable and predictive tool in measuring noise tolerance while maintaining subjective speech intelligibility. The similar TNL estimated from the aTNL and eTNL scoring methods suggests that the TNL tracking duration can be shortened for many participants ($>$65%). In addition, the TNT procedure revealed a within-session reliability of 1.5 dB using babble noise and 2.8 dB using continuous speech-shaped noise. Finally, the TNL showed a significant relationship between the amount of noise tolerated and the hearing aid wearers' subjective satisfaction in loud noisy environments. Additional data collection with more participants covering a greater variety of satisfaction levels (in loud noise) will enhance our understanding of such potential relationship.

Acknowledgments. We extend our gratitude and appreciation to the external sites who participated in data collection. These sites included Marco Hearing Health Centre and Davidson Hearing Aid Centre, both located in Canada. Furthermore, we are thankful for the test participants' time and effort.

REFERENCES

American Academy of Audiology (AAA), Task Force for Guidelines for the Audiologic Management of Adult Hearing Impairment. (2006) *Guidelines for the Audiologic Management of Adult Hearing Impairment.* Retrieved from https://audiology-web.s3.amazonaws.com/migrated/haguidelines.pdf_53994876e92e42.70908344.pdf. Accessed June 1, 2017.

American National Standards Institute (ANSI) (1997) American National Standard: Methods for the Calculation of the Speech Intelligibility Index. ANSI S3.5. New York, NY: ANSI.

Beattie RC, Boyd RL. (1986) Relationship between pure tone and speech loudness discomfort levels among hearing-impaired subjects. *J Speech Hear Disord* 51(2):120–124.

Bentler R, Wu YH, Kettel J, Hurtig R. (2008) Digital noise reduction: outcomes from laboratory and field studies. *Int J Audiol* 47(8):447–460.

Bertoli S, Staehelin K, Zemp E, Schindler C, Bodmer D, Probst R. (2009) Survey on hearing aid use and satisfaction in Switzerland and their determinants. *Int J Audiol* 48(4):183–195.

Brännström J, Olsen S, Holm L, Kastberg T, Ibertsson T. (2014) The effect of repeated measurements and working memory on the most comfortable level in the ANL test. *Int J Audiol* 53(11):787–795.

Brungart DS. (2001) Informational and energetic masking effects in the perception of two simultaneous talkers. *J Acoust Soc Am* 109(3):1101–1109.

Carhart R. (1947) Selection of hearing aids. *Arch Otolaryngol* 44: 1–18.

Cooke M. (2006) A glimpsing model of speech perception in noise. *J Acoust Soc Am* 119(3):1562–1573.

Cox RM, Alexander GC. (1995) The abbreviated profile of hearing aid benefit. *Ear Hear* 16(2):176–186.

Cox RM, Alexander GC. (2002) The International Outcome Inventory for Hearing Aids (IOI-HA): psychometric properties of the English version. *Int J Audiol* 41(1):30–35.

Davis H, Hudgins C, Marquis R, Nichols R, Peterson G, Ross D, Stevens S. (1946) The selection of hearing aids. *Laryngoscope* 56:85–115.

Dawes P, Maslin M, Munro KJ. (2014) 'Getting used to' hearing aids from the perspective of adult hearing aid users. *Int J Audiol* 53(12):861–870.

Dillon H. (2001) *Hearing Aids*. New York, NY: Thieme.

Filion P, Margolis RH. (1992) Comparison of clinical and real-life judgments of loudness discomfort. *J Am Acad Audiol* 3(3):193–199.

Franklin CA, White LJ, Franklin TC. (2012) Relationship between loudness tolerance and the acceptance of background noise for young adults with normal hearing. *Percept Mot Skills* 114(3):717–722.

Freyaldenhoven MC, Nabelek AK, Burchfield SB, Thelin JW. (2005) Acceptable noise level as a measure of directional hearing aid benefit. *J Am Acad Audiol* 16(4):228–236.

Freyaldenhoven MC, Nabelek AK, Tampas JW. (2008) Relationship between acceptable noise level and the abbreviated profile of hearing aid benefit. *J Speech Lang Hear Res* 51(1):136–146.

Freyaldenhoven MC, Smiley DF, Muenchen RA, Knorad TN. (2006) Acceptable noise level: reliability measures and comparison to Preference for background sounds. *J Am Acad Audiol* 17(9):640–648.

Hawthorne G. (2008) Perceived social isolation in a community sample: its prevalence and correlates with aspects of peoples' lives. *Soc Psychiatry Psychiatr Epidemiol* 43(2):140–150.

Ho HC, Wu YH, Hsiao SH, Zhang X. (2013) Acceptable noise level (ANL) and real-world hearing-aid success in Taiwanese listeners. *Int J Audiol* 52(11):762–770.

Jenstad LM, Van Tasell DJ, Ewert C. (2003) Hearing aid troubleshooting based on patients' descriptions. *J Am Acad Audiol* 14(7):347–360.

Keidser G, Dillon H, Flax M, Ching T, Brewer S. (2011) The NAL-NL2 prescription procedure. *Audiol Res* 1(1):e24.

Kincaid JP, Fishburne RP, Rogers RL, Chissom BS. (1975) Derivation of new readability formulas: (automated readability index, fog count, and Flesch reading ease formula) for Navy enlisted personnel [Research Branch Report 8–75]. Chief of Naval Technical Training: Naval Air Station Memphis.

Kochkin S. (2000) MarkeTrak V: 'why my hearing aids are in the drawer': the consumers' perspective. *Hear J* 53:34–41.

Kochkin S. (2010) MarkeTrak VIII: consumer satisfaction with hearing aids is slowly improving. *Hear J* 63(1):19–32.

Krishnamurti S, Anderson L. (2008) Digital noise reduction processing in hearing aids: how much and where? *Hear Rev* 15(3):90–93.

Kuk F, Damsgaard A, Bulow M, Ludvigsen C. (2004) Using digital hearing aids to visualize real life effects of signal processing. *Hear J* 57(4):40–49.

Kuk F, Lau CC, Korhonen P, Crose B. (2015) Speech intelligibility benefits of hearing aids at various input levels. *J Am Acad Audiol* 26(3):275–288.

Kuk F, Peeters H, Lau C, Korhonen P. (2011) Effect of maximum power output and noise reduction on speech recognition in noise. *J Am Acad Audiol* 22(5):265–273.

Kuk F, Seper E, Lau C, Korhonen P. (2017) Tracking of noise tolerance to measure hearing aid benefit. *J Am Acad Audiol* 28(8):698–707.

Morgan DE, Wilson RH, Dirks DD. (1974) Loudness discomfort level: selected methods and stimuli. *J Acoust Soc Am* 56(2): 577–581.

Mueller HG. (2003) Fitting test protocols are more honored in the breach than in the observance. *Hear J* 56(10):19–26.

Mueller HG, Bentler RA. (2005) Fitting hearing aids using clinical measures of loudness discomfort levels: an evidence-based review of effectiveness. *J Am Acad Audiol* 16(7):461–472.

Nabelek AK, Freyaldenhoven MC, Tampas JW, Burchfield SB, Muenchen RA. (2006) Acceptable noise level as a predictor of hearing aid use. *J Am Acad Audiol* 17(9):626–639.

Nabelek AK, Tucker FM, Letowski TR. (1991) Toleration of background noise: relationship with patterns of hearing aid use by elderly persons. *J Speech Hear Res* 34:679–685.

Oeding K, Valente M. (2013) Differences in sensation level between the Widex soundtracker and two real-ear analyzers. *J Am Acad Audiol* 24(8):660–670.

Olsen SO, Brännström KJ. (2014) Does the acceptable noise level (ANL) predict hearing aid use? *Int J Audiol* 53(1):2–20.

Olsen SO, Lantz J, Nielsen LH, Brännström KJ. (2012) Acceptable noise level (ANL) with Danish and non-semantic speech materials in adult hearing-aid users. *Int J Audiol* 51(9):678–688.

Olsen S, Nielsen L, Lantz J, Brännström J. (2013) Intertester reliability of the acceptable noise level. *J Am Acad Audiol* 24(3): 241–243.

Pearsons K, Bennett R, Fidell S. (1977) *Speech Levels in Various Noise Environments* [Report No. EPA-600/1-77-025]. Washington, DC: US Environmental protection Agency.

Peeters H, Kuk F, Lau CC, Keenan D. (2009) Subjective and objective evaluation of noise management algorithms. *J Am Acad Audiol* 20(2):89–98.

Philibert B, Collet L, Vesson JF, Veuillet E. (2005) The auditory acclimatization effect in sensorineural hearing-impaired listeners: evidence for functional plasticity. *Hear Res* 205(1):131–142.

Stephens SDG, Anderson CMB. (1971) Experimental studies on the uncomfortable loudness level. *J Speech Hear Res* 14: 262–270.

Taylor B. (2008) The acceptable noise level test as a predictor of real-world hearing aid benefit. *Hear J* 61(9):39–42.

Walravens E, Keidser G, Hartley D, Hickson L. (2014) An Australian version of the acceptable noise level test and its predictive value for successful hearing aid use in an older population. *Int J Audiol* 53:S52–S59.

Watson L, Tolan T. (1949) *Hearing Tests and Hearing Instruments*. Baltimore, MD: Williams & Wilkins.

Wu YH, Ho HC, Hsiao SH, Brummet RB, Chipara O. (2016) Predicting three-month and 12-month post-fitting real-world hearing-aid outcome using pre-fitting acceptable noise level (ANL). *Int J Audiol* 55(5):285–294.

Wu YH, Stangl E. (2013) The effect of hearing aid signal-processing schemes on acceptable noise levels: perception and prediction. *Ear Hear* 34(3):333–341.