Acceptable noise level: Repeatability with Danish and non-semantic speech materials for adults with normal hearing

Olsen, Steen Ostergaard; Nielsen, Lars Holme; Lantz, Johannes; Brännström, Jonas

Published in:
International Journal of Audiology

DOI:
10.3109/14992027.2012.666362

2012

The Acceptable Noise Level: Repeatability with Danish and non-semantic speech materials for adults with normal hearing

Steen Østergaard Olsen[ad], Lars Holme Nielsen[a], Johannes Lantz[bd], K. Jonas Brännström[bcd]

[a]Research Laboratory, Department of Otorhinolaryngology, Head and Neck Surgery, University Hospital, Rigshospitalet, Copenhagen, Denmark.

[b]Section of Audiology, ENT-department, Skåne University Hospital in Malmö, Sweden.

[c]Department of Clinical Science, Section of Logopedics, Phoniatrics and Audiology, Lund University, Lund, Sweden.

[d]HEARsound Laboratories, Network for Joint Hearing Research in the Oresunds Region.

Corresponding author: Steen Østergaard Olsen

Address: Research Laboratory

Department of Otorhinolaryngology, Head and Neck Surgery

Blegdamsvej 9

DK-2100 Copenhagen, Denmark

Phone: +45 35 45 22 79

Fax: +45 35 45 26 29

E-mail: steen.olsen@rh.regionh.dk

Abstract

Objective: The acceptable noise level (ANL) is used to quantify the amount of background noise that subjects can accept while listening to speech, and is suggested for prediction of individual hearing aid use. The aim of this study was to assess the repeatability of the ANL measured in normal-hearing subjects using running Danish and non-semantic speech materials as stimuli and modulated speech-spectrum and multi talker babble noises as competing stimuli. **Design:** ANL was measured in both ears at two test sessions separated by a period ranging from 12 to 77 days. At each session the measurements at the first and the second ear were separated in time by 15-30 minutes. Bland-Altman plots and calculation of the coefficient of repeatability (CR) were used to estimate the repeatability. **Study Sample:** Thirty nine normal-hearing subjects. **Results:** The ANL CR was 6.0 – 8.9 dB for repeated tests separated by about 15-30 minutes and 7.2 – 10.2 dB for repeated tests separated by 12 days or more. **Conclusions:** The ANL test has poor repeatability when assessed with Danish and non-semantic speech materials on normal-hearing subjects. The same CR among hearing impaired subjects would imply too poor repeatability to predict individual patterns of future hearing aid use.

Key words: Acceptable Noise Level, Danish, test-retest, International Speech Test Signal, normal-hearing, repeatability.

Abbreviations: ANL = Acceptable Noise Level, ANOVA = Analysis of variance, BNL = Background Noise Level, CR = Coefficient of repeatability, MCID = minimal clinically important difference, ILTASS = International Long-term Average Speech Spectrum, ISTS = International Speech Test Signal, MCL = Most Comfortable Level, PTA = Pure Tone Average, RMS = Root-Mean-Square, SD = standard deviation, SPL = Sound Pressure Level

Introduction

Measurement of the acceptable noise level (ANL) was introduced by Nabelek et al (1991) as a method for quantification of the amount of background noise that subjects can accept when listening to speech at the most comfortable level (MCL). The ANL in dB is the difference between the MCL for running speech and the highest background noise level (BNL) that is acceptable when following the running speech (Nabelek et al, 2004). ANL has been used for prediction of individual hearing aid use patterns (full time use or occasional use/no use at all), and has also been used to assess the benefit of hearing aid algorithms for reduction of background noise (Freyaldenhoven et al, 2005a; Mueller et al, 2006). Danish and Swedish versions of the ANL test were described by Brännström et al (2011).

Ideally, a psychoacoustic test like ANL should be repeatable, that is, the test should yield similar results when measured on the same subject on different occasions. Earlier studies of normal-hearing and hearing-impaired adult subjects have shown a high correlation between ANL measurements obtained at three different sessions (Nabelek et al, 2004; Freyaldenhoven et al, 2006). In the study by Nabelek et al (2004) the time between the first and the second session was one month and between the second and the third session it was two months. In the study by Freyaldenhoven et al (2006) the intersession time was approximately 1 week.

The correlation coefficient shows the association, not the agreement between repeated measures. Two measures (x and y) are associated with each other if for instance the relationship between the measures is $y = 0.5x$, but two measures only agree if the relation is $y = x$. To assess the repeatability of a method it is necessary to study the agreement of the measures and according to Bland and Altman (1986), the use of the correlation coefficient is inappropriate for estimation of the repeatability. In the present study we therefore calculated the coefficient of repeatability (CR), which can show a potential agreement between repeated measures (Bland and Altman, 1986). Provided that the data are normally distributed, the CR is the value below which the absolute difference between two repeated test results may be expected to lie with a probability of 95%. The CR is calculated as 1.96 x the standard deviation (SD) of the differences.

The aim of the present study was to investigate the repeatability among normal-hearing Danish subjects when listening to two types of running speech (Danish and a speech signal without semantic content) in two types of competing noise (modulated speech shaped noise and speech babble).

Material and methods

Fifty-five Danish subjects were initially recruited to the study. Fifteen subjects did not meet the inclusion criteria (see below) and were excluded. One of the subjects did only show up for the first of two sessions and thus thirty nine adults (24 female and 15 male) participated in the study. The native language of the subjects was Danish. All subjects were also included in an earlier study (Brännström et al, 2011), and the raw data analyzed in the present study are the same. The participants (median age: 31 years; range: 19–54 years) were university students, colleagues, family and friends of the authors. The inclusion criteria in the study were described by Brännström et al (2011) and consisted of normal tympanograms and recordable acoustic reflexes at 1 kHz and pure-tone hearing thresholds $\leq$ 15 dB HL for octave frequencies 250 to 4000 Hz. The subjects' demographic data with an audiometric summary are presented in table 1. None of the included subjects had previously performed the ANL test. The subjects received a symbolic gift for their participation. The Scientific Ethical Review Board C for the Capital Region approved the project (H-C-2009-022.)

Speech and noise signals

The ANL stimuli consisted of a speech signal in one channel and a noise signal in the other channel, which were presented monaurally to the same earphone. These signals were pre-recorded and implemented in the audiometer software as integrated speech stimulus files. Three test conditions were used: (1) running speech and speech-weighted amplitude-modulated noise, (2) running speech and multi talker speech babble, and (3) International Speech Test Signal (ISTS) (Holube et al, 2010) and speech-weighted amplitude-modulated noise.

The Dantale audio compact disc (Elberling et al, 1989) is available on all Danish hearing clinics. Therefore the running speech signal from this disc (track 12) was used as speech signal in conditions 1 and 2. The track contains a geographic and historical description (Andersen, 1983) read by a female speaker. The duration of the story is 4 minutes and 23 seconds and contains on average about 40 syllables per 10 seconds. Also, approximately 16% of the total duration is silence (i.e. pauses). The signal is low-pass filtered at 10 000 Hz. In condition 3 the ISTS was used as speech signal. The ISTS contains concatenated syllables from six selected female speakers of six different languages (Arabic, English, French, German, Mandarin, and Spanish) reading the story “The North Wind and the Sun”. The recordings were first filtered to match the International Long-term Average Speech Spectrum (ILTASS) (Byrne et al, 1994) and then split into segments, roughly approximating one syllable. Segments from the different languages were then concatenated into utterances closely resembling running speech according to strict rules. The resulting signal sounds like speech spoken in an unfamiliar language. Brännström et al (2011) reported a smaller inter-individual SD for ANL when using ISTS as speech signal in speech-weighted noise, and suggested that this might improve the reliability of the test.

The noise from the Dantale audio compact disc (track 12, Dantale noise) was used as noise signal in conditions 1 and 3. The Dantale noise and its recording have been described in detail by Elberling et al (1989). This noise is speech-spectrum shaped and generated from band-pass filtered white noise. The noise is amplitude-modulated using a band-pass filtered white noise at 4 Hz and has a modulation depth of approximately 27 %. In the following the Dantale noise is referred to as speech-weighted noise. In condition 2, the ANL multi talker (12 voices) speech-babble noise was used. This noise was taken from the official ANL CD (Arizona Travelodge, Cosmos Distributing Inc.) and is identical to the noise used in the Revised Speech in Noise Test by Bilger et al (1984).

The signals were extracted from their digital sources and assembled using the computer software Adobe Audition (version 3.0). A sampling rate of 44 100 Hz was used. The average RMS-values for these signals were equalised and all showed levels of about -20 dB relative to the point of peak clipping, i.e. the risk of distortion due to peak clipping was excluded.

Equipment and calibration for pure-tone audiometry and ANL

The complete equipment set-up (Astera Audiometer and the HDA 200 earphones) was calibrated before the study started, in accordance with IEC 60318-2 (1998), ISO 389-5 (2006), and ISO 389-8 (2004) using the Brüel and Kjaer 2610 measuring amplifier with a 4144 microphone in a 4152 ear simulator. The frequency-modulated 1000 Hz calibration tone (modulation rate = 20 Hz, modulation frequency = 250 Hz) from the Dantale CD with a 10 dB higher average RMS than the speech and noise signals was used for the calibration of these signals. The actual output levels of the speech and noise signals presented through the transducers were also verified and all signals showed the same average dB SPL RMS ± 1 dB (averaging time was 20 seconds, linear frequency weight) at identical audiometer output settings. All tests were performed in a double-walled soundproof booth (complying with the maximum permissible ambient sound pressure levels as specified in ISO 8253-1 (1998)) during two sessions.

Procedures

The first session started with a short interview to ascertain that the subject was otologically healthy. Otoscopy and immittance measurements were followed by pure-tone audiometry and finally, ANL measurements were made. Both written and oral instructions were given prior to ANL testing. The instructions (Brännström et al, 2011) were Danish versions of the English instructions (Nabelek et al, 1991). If the subjects had any doubts the instructions were clarified. Examples of speech and noise were then presented. The subjects were instructed to listen to the speech signal in quiet presented monaurally through the earphones and, after a few moments of listening, to adjust the sound level to the MCL using an up-and-down procedure; the subject used the audiometer attenuators (without any visual feedback) to increase the loudness until the speech signal became too loud, then decreased it until it became too soft, and finally the subject selected the loudness that was found most comfortable (MCL). The subject reported when the MCL was found. Then BNL was established by adding a noise signal in the same earphone as the speech signal and the subject was instructed to repeat a similar procedure; the speech signal remained fixed at the previously established MCL and the subject increased the loudness of the noise until it became too loud, then decreased it until the speech became very clear, and finally the subject selected the loudness (i.e. BNL) that the subject could accept or “put up with” without becoming tense or tired while following the speech signal for a long period of time. The subject reported when the BNL had been found. A 2 dB step size was used for all adjustments for both MCL and BNL. The complete procedure (for both MCL and BNL) was repeated three times, and the outcome was calculated as the mean of the three measurements. No time limit was set for establishing either the MCL or the BNL, but all measurements on one ear (three repetitions of three conditions) took about 15 to 30 minutes depending on the individual subject. When all the measurements had been done on one ear the subjects were offered a small break before proceeding with the second ear, but all subjects with very few exceptions preferred to continue without a break. The initial stimulus level for each repetition was 58 dB SPL for both speech and noise signals. The test condition presentation order and test ear order was balanced across subjects according to a Latin squares design.

At the start of the second session the written and oral instructions were repeated. Then the ANL measurements were carried out exactly as at the first session.

All ANL measurements at the two sessions were carried out by the same experimenter.

Statistical analysis

We estimate that the minimal clinically important difference (MCID) when measuring the ANL is about 4 dB. The MCID is not a fixed value but has to be selected according to the context (Beaton et al, 2002). In this paper we use “The smallest difference in a score that is considered to be important” as the definition for MCID. To determine the MCID we used the following opinion-based approach: ANL is used to predict future hearing aid use patterns, but also to evaluate benefit from hearing aid algorithms. The MCID should have a size which is sufficient to let the user of a certain hearing instrument feel the benefit from an algorithm designed to reduce the amplification of noise in the environment. An improvement of the signal-to-noise ratio of about 3-4 dB due to the use of directional microphones is often reported as a benefit for hearing users (e.g. Wouters et al, 1999). As a consequence we selected 4 dB as the MCID for ANL.

In eight ANL studies with normal subjects (Nabelek et al, 1991; Rogers et al, 2003; Franklin et al, 2006; Freyaldenhoven et al, 2006; Tampas and Harkrider, 2006; von Hapsburg and Bahng, 2006; Freyaldenhoven et al, 2007) a wide range of inter-individual SD's were reported. The median inter-individual SD across the studies was 5.3 dB (range: 1.8-11.0). With an inter-individual SD of 5 dB and a MCID of 4 dB one would achieve a statistical power of around 90% at the 5% significance level (Altman, 1991). The statistical power is the probability of not committing a Type II error. An alpha level < 0.05 was considered statistically significant throughout the study.

According to Bland and Altman (1986), the CR should be less than the MCID. For the ANL test the CR should therefore be less than 4 dB.

Differences between the ears in MCL, BNL, and ANL were explored using a paired-samples T-test (alpha level, $p<0.05$). Gender differences were explored using two independent-samples T-test; the same alpha level was used as for the paired-samples test. Association between the collected variables and demographic data (such as age and PTA) were explored using Pearson's ρ; probability values, p, <0.05 were considered statistically significant.

Repeated measures analyses of variance (ANOVA) were used to assess the effects of test ear order (first and second ear) and test session (first and second session) for MCL, BNL, and ANL within each test condition; alpha levels < 0.05 were considered statistically significant. In each analysis, one within-subject variable (test ear order or test session) was used. Single measure intraclass correlation coefficients were calculated to investigate possible associations between the outcomes obtained at the two sessions.

Parametric tests were used since the average MCL, BNL, and ANL results in each condition in both sessions were found to be normally distributed using the Kolmogorov-Smirnov test for normality.

Results

The median period between the first and the second session was 15 days (range 12-77 days).

Differences and associations

Associations within subjects was assessed using repeated measures ANOVA examining the within-subject main effects of test session for MCL, BNL, and ANL in each test condition and ear. Overall, no significant interaction effects were observed. No significant main effects of test session were seen for MCL, BNL, or ANL test repetition. Similar SDs and ranges for MCLs and BNLs were seen in all conditions for both test sessions. Overall, the findings suggested no effect of test session.

No significant gender differences were found for MCL, BNL, or ANL in any ear. No significant differences between right and left ear were observed for MCL, BNL, or ANL within each session. No associations were observed between age, PTA in left or right ear, and MCL, BNL, or ANL results. A significant association was seen between age and PTA in the right ear only (DF =10, $\rho = 0.427$, $p = 0.007$).

Figure 1 depicts Bland-Altman plots showing differences between ANL measurements at the first and the second ear separated in time by approximately 15-30 minutes. The differences are plotted as a function of the mean of the two ANL measurements and are shown for the three conditions. No relation between the spread of the differences and the size of the mean ANL is seen. The mean difference across subjects lies in all three conditions at the two sessions between -1.1 and -1.9 dB (Table 2) showing an ANL increase from the first to the last measured ear. The CR in all three conditions lies between 6.0 – 8.9 dB (Table 2). There is no apparent trend indicating that one condition has a better repeatability (lower CR) than others.

Table 2 shows the Mean (M), standard deviation (SD), and ranges of differences between the MCL, BNL, and ANL measured at the first and the last measured ear at Session 1 and 2 for all subjects in the three test conditions. An asterisk indicates that the mean differences between ANL measurements are statistically significant. Also shown are the ANL coefficients of repeatability (CR). Measurements were separated in time by approximately 15-30 minutes.

Figure 2 depicts Bland-Altman plots showing the differences between ANL measurements at the first and the last session as a function of the mean ANL obtained at the two sessions. Plots are shown for the three conditions for each ear separately. No relation between the spread of the differences and the size of the mean ANL is seen. The mean of the differences on both ears in all three conditions are close to zero (0.1 to 0.8 dB). The CR in all three conditions lies between 7.2 – 10.2 dB (Table 3). There is no apparent trend indicating that one condition has a better repeatability than others.

Table 3 shows the Mean (M), standard deviation (SD), and ranges of differences between the MCL, BNL, and ANL measured at session 1 and 2 for all subjects in the three test conditions. Also shown are the ANL coefficients of repeatability (CR). Measurements were separated in time by at least 12 days.

To examine the within-subject main effects of ear and session order for MCL, BNL, and ANL in each test condition a repeated measures ANOVA was used. Overall, no significant interaction effects were observed. No significant main effects of test order were seen for MCL and BNL. Significant main effects of test order were seen for ANL in all test conditions ($F > [4.15]$, $p < 0.05$) except for condition 2 although approaching significance ($F[3.91]$, $p = 0.055$). This indicates that ANL increases (becomes poorer) in the second ear in all test conditions. No significant main effects of test session were seen for MCL, BNL, or ANL.

Discussion

An earlier study of normal hearing subjects (Freyaldenhoven et al, 2006) showed a high correlation between ANL measurements done at different occasions. The correlation coefficient shows the association of repeated measures, but it does not show the agreement between measures (Bland and Altman, 1986). In this study we calculated the CR which is capable of showing a potential agreement between repeated measures (Bland and Altman, 1986). Calculation of the SD of differences of repeated measurements has been used for the assessment of the repeatability of pure tone audiometry (Jerlvall and Arlinger, 1986), but we have not been able to identify previous ANL-studies in which the SD or Bland and Altmans (1986) method for repeatability assessment were used. In a letter to the editor, Nabelek et al (2007) reported that in a study of normal-hearing subjects (Freyaldenhoven et al, 2006), individual test-retest ANL differences were up to about 14 dB, which is comparable to our results (table 2 and 4). These individual test-retest differences are not reported in the original paper (Freyaldenhoven et al, 2006).

Freyaldenhoven et al (2005b) reported that medication significantly increased the acceptance of background noise for individuals with ADHD/ADD. The subjects in our study were not inquired whether or not they were in treatment with centrally stimulating or depressing medications. Medication might in individual cases have affected the outcome of the measurements.

Previous studies concluded that ANL is an individual inherent feature, unaffected by factors like age and degree of hearing loss (Nabelek et al, 2004; Nabelek et al, 2007). In the present study no systematic differences were seen between the ANL on right and left ear. We therefore found it acceptable to compare measurements on the first measured ear with measurements on the last measured ear to assess the short term ANL repeatability.

According to Nabelek et al (2006) subjects with an ANL less than 7 dB are likely to become successful, full-time hearing aid users, while subjects with an ANL more than 13 dB are likely to become unsuccessful users who wear hearing aids occasionally or not at all. For subjects with an ANL between 7 and 13 dB the use pattern cannot be predicted. The ANL CR in the present study was 6.0 – 8.9 dB, when the two measurements were separated in time by approximately 15-30 minutes, and 7.2 – 10.2 dB, when the two measurements were separated in time by more than twelve days. Our results are from normally hearing subjects and a similar range of CRs may not be found in hearing impaired subjects. On the other hand, should that be the case, then hypothetically one single subject might be categorized as a potential full-time hearing aid user based on one measurement (ANL < 7 dB), and after another measurement be expected to use hearing aid only occasionally or not at all (ANL > 13 dB). That would imply that the ANL test has too poor repeatability to predict individual patterns of future hearing aid use. In the already mentioned letter to the editor, Nabelek et al (2007) also reported that in a study (Nabelek et al, 2004), the test-retest difference for individual hearing-impaired listeners was approximately 2 dB. This value was not directly reported in the original paper (Nabelek et al, 2004), but figure 1 in the paper shows test-retest differences of about 0-4 dB. These differences are much smaller than the individual differences found in the present study, and it seems that this paper would have demonstrated a strong CR. However, the present study examines normal-hearing individuals and there might be differences in psychometric functions for normal-hearing and hearing-impaired subjects. Further studies are required to examine this issue. Why the test-retest difference was as small as 2 dB in the study by Nabelek et al (2004) with hearing-impaired subjects and much higher in the present study with normal-hearing subjects is unclear. In the study by Nabelek (2004) the underlying pathology causing the hearing losses found among the subjects was not described. If, however, the fifty subjects in the study hearing impaired had sensorineural losses with recruitment, the reduced dynamic range might be a contributing factor to the smaller test-retest differences. In the present study all the subjects did the ANL test for the first time ever at the first session. In the study by Nabelek et al (2004) the subjects were also included in an earlier study, but it is not clear whether the participants underwent new ANL measurements for the study or if the data from earlier studies were just reanalyzed. Participation in several ANL studies might result in a better repeatability.

Our results show a small increase in ANL from the first to the last tested ear, when the measurements are separated in time by approximately 15-30 minutes. An increase in ANL means that the listeners in our study accept less competing noise, when they have been doing the test for a while. This might be caused by fatigue or by some other order effect. Investigation of such effects was not the scope of the present study, and we are not aware of any reports from studies on the issue, but it would be valuable to investigate these putative effects.

Conclusion

The repeatability of the ANL test with Danish and non-semantic speech materials is poor when measured on normal-hearing subjects. If the coefficient of repeatability is of the same magnitude among hearing impaired subjects, the repeatability would be too poor to predict individual patterns of future hearing aid use. Further studies are required.

Acknowledgements

The authors thank engineer Arne Månsson for valuable discussions and advices regarding equipment calibration and Anders H Rasmussen for integrating the test signals in the audiometric software.

Declaration of interest

No outside funding or grants in support of this research were received. The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

References

Altman D.G. 1991. *Practical statistics for medical research*. London: Chapman & Hall.

Andersen, J.R. 1983. *Samsø rundt*. Tranebjerg: Flemming Andersens Bookstore.

Beaton, D.E., Boers M., Wells G.A. 2002. Many faces of the minimal clinically important difference (MCID): a literature review and directions for future research. *Curr Opin Rheumatol*, 14, 109–114.

Bilger, R.C., Nuetzel, J.M., Rabinowitz, W.M., Rzeczkowski, C. 1984. Standardization of a test of speech perception in noise. *J Speech Hear Res*, 27, 32-48.

Bland, J.M., Altman, D.G. 1986. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*, 327, 307–310.

Brännström, K.J., Lantz, J., Nielsen, L.H., Olsen. S.Ø. (2011) Acceptable Noise Level with Danish, Swedish and non-semantic speech materials. *Int J Audiol*, 50, Oct 24. [Epub ahead of print].

Byrne, D., Dillon, H., Tran, K., Arlinger, S., Wilbraham, K. et al. 1994. An international comparison of long-term average speech spectra. *J Acoust Soc Am*, 96, 2108-2120.

Elberling, C., Ludvigsen, C., Lyregaard, P.E. 1989. Dantale: a new Danish speech material. *Scand Audiol*, 18, 169-175.

Franklin, C.A., Thelin, J.W., Nabelek, A.K., Burchfield, S.B. 2006. The effect of speech presentation level on acceptance of background noise in listeners with normal hearing. *J Am Acad Audiol*, 17, 141–146.

Freyaldenhoven, M.C., Nabelek, A.K., Burchfield, S.B., Thelin, J.W. (2005a). Acceptable noise level as a measure of directional benefit. *J Am Acad Audiol*, 16, 228–236.

Freyaldenhoven, M.C., Plyler, P.N., Thelin, J.W., Hedrick, M.S. 2007. The effects of speech presentation level on acceptance of noise in listeners with normal and impaired hearing. *J Speech Lang Hear Res*, 50, 878-885.

Freyaldenhoven, M.C., Smiley, D.F., Muenchen, R.A., Konrad, T.N. 2006. Acceptable noise level: reliability measures and comparison to preference for background sounds. *J Am Acad Audiol*, 17, 640-648.

Freyaldenhoven, M. F., Thelin, J. W., Plyler, P. N., Nabelek, A. K., Burchfield, S. B. 2005b. Effect of stimulant medication on the acceptance of background noise in individuals with attention deficit/hyperactivity disorder. *J Am Acad Audiol*, 16, 677–686.

Holube, I., Fredelake, S., Vlaming, M., Kollmeier, B. 2010. Development and analysis of an International Speech Test Signal (ISTS). *Int J Audiol*, 49, 891-903.

IEC 60318-2. 1998. Electroacoustics - Simulators of human head and ear - Part 2: An interim acoustic coupler for the calibration of audiometric earphones in the extended high-frequency range International Electrotechnical Commission. Geneva.

ISO 389-5. 2006. Acoustics: Reference zero for the calibration of audiometric equipment. Part 5: Reference equivalent threshold sound pressure levels for pure tones in the frequency range 8 kHz to 16 kHz. International Organization for Standardization 389-5.

ISO 389-8. 2004. Acoustics: Reference zero for the calibration of audiometric equipment. Part 8: Reference equivalent threshold sound pressure levels for pure tones and circumaural earphones. International Organization for Standardization 389-8.

ISO 8253-1. 1998. Acoustics: Audiometric test methods part 1: Basic pure tone air and bone conduction threshold audiometry. International Organization for Standardization 8253-1.

Jerlvall, L., Arlinger, S. 1986. A comparison of 2-dB and 5-dB step size in pure-tone audiometry. *Scand Aud*, 15, 51-56.

Mueller, H.G., Weber, J., Hornsby, B.W. 2006. The effects of digital noise reduction on the acceptance of background noise. *Trends Ampl*, 10, 83-93.

Nabelek, A.K., Freyaldenhoven, M.C., Tampas, J.W., Burchfiel, S.B., Muenchen, R.A. 2006. Acceptable noise level as a predictor of hearing aid use. *J Am Acad Audiol*, 17, 626-639.

Nabelek, A.K., Tampas, J.W., Burchfield, S.B. 2004. Comparison of speech perception in background noise with acceptance of background noise in aided and unaided conditions. *J Speech Lang Hear Res*, 47, 1001-1011.

Nabelek, A.K., Tampas, J.W., Freyaldenhoven, M.C. 2007. Further questions about the acceptable noise level test: A response to Dr. Hamill. *J Am Acad Audiol*, 18, 185-187.

Nabelek, A.K., Tucker, F.M., Letowski, T.R. 1991. Toleration of background noises: relationship with patterns of hearing aid use by elderly persons. *J Speech Hear Res*, 34, 679-685.

Rogers, D.S., Harkrider, A.W., Burchfield, S.B., Nabelek, A.K. 2003. The influence of listener's gender on the acceptance of background noise. *J Am Acad Audiol*, 14, 372-382.

Tampas, J.W., Harkrider, A.W. 2006. Auditory evoked potentials in females with high and low acceptance of background noise when listening to speech. *J Acoust Soc Am*, 119, 1548–1561.

von Hapsburg, D., Bahng, J. 2006. Acceptance of background noise levels in bilingual (Korean-English) listeners. *J Am Acad Audiol*, 17, 649–658.

Wouters, J., Litière, L., van Wieringen, A. 1999. Speech intelligibility in noisy environments with one- and two-microphone hearing aids. *Audiology*, 38, 91-98.

Table 1. **The subjects' demographic data with audiometric summary. PTA was calculated as the average in dB HL for frequencies 0.5, 1, 2, and 4 kHz.**

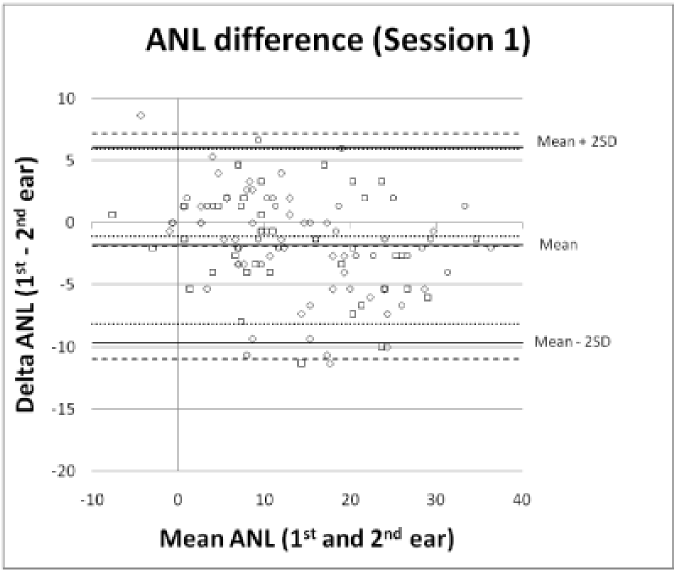
Variable			**Subjects (n=39)**	
Age (years)		Mean	33.4	
		SD	9.7	
		Range	19-54	
Gender		Female	61.5 %	(n=24)
		Male	38.5 %	(n=15)
PTA (dB HL)	Right ear	Mean	2.6	
		SD	5.0	
		Range	-10 - 15	
	Left ear	Mean	2.2	
		SD	5.2	
		Range	-10 - 15	

Table 2. **Mean (M), standard deviation (SD), and ranges of MCL, BNL, ANL differences and CR of the first and the last measured ear. An asterisk indicates that the mean differences between ANL measurements are statistically significant.**

			First ear - Second ear (n=39)	
			Session 1	**Session 2**
Condition 1: Running speech and Dantale noise	**MCL**	M	-0.5	-0.2
		SD	5.1	2.7
		Range	-8.7-13.3	-6.0-4.7
	BNL	M	1.3	1.5
		SD	5.1	3.9
		Range	-7.3-12.0	-7.3-10.0
	ANL	M	-1.8*	-1.7*
		SD	3.9	3.7
		Range	-11.3-4.7	-8.7-6.7
		CR	7.7	7.3
Condition 2: Running speech and ANL babble	**MCL**	M	-1.2	-1.0
		SD	4.5	3.3
		Range	-19.3-7.3	-10.0-6.0
	BNL	M	-0.1	0.0
		SD	4.6	3.3
		Range	-10.7-8.7	-7.3-6.7
	ANL	M	-1.1	-1.0*
		SD	3.5	3.1
		Range	-10.7-6.7	-8.7-4.7
		CR	6.9	6.0
Condition 3: ISTS as speech and Dantale noise	**MCL**	M	-0.6	-0.2
		SD	4.1	2.9
		Range	-11.3-8.0	-5.3-6.7
	BNL	M	1.3	1.1
		SD	5.3	4.2
		Range	-11.3-10.7	-14.0-8.0
	ANL	M	-1.9*	-1.3*
		SD	4.5	3.7
		Range	-11.3-10.7	-10.0-12.0
		CR	8.9	7.2

Table 3. **Same as table 2, but showing differences between repeated measurements with an interval of 12 days or more.**

			Session 1-Session 2 (n=39)	
			Right ear	**Left ear**
Condition 1: Running speech and Dantale noise	**MCL**	M	0.7	0.9
		SD	3.8	4.4
		Range	-6.7-11.3	-6.7-11.3
	BNL	M	0.1	0.5
		SD	5.3	6.3
		Range	-10.7-12.0	-11.3-24.0
	ANL	M	0.5	0.4
		SD	5.2	5.0
		Range	-11.3-12.0	-12.7-14.0
		CR	10.2	9.7
Condition 2: Running speech and ANL babble	**MCL**	M	0.2	0.7
		SD	4.4	3.6
		Range	-13.3-10.7	-8.0-10.0
	BNL	M	-0.6	0.1
		SD	5.5	5.4
		Range	-13.3-16.7	-8.0-14.7
	ANL	M	0.8	0.6
		SD	5.1	5.0
		Range	-14.0-12.0	-11.3-11.3
		CR	10.0	9.7
Condition 3: ISTS as speech and Dantale noise	**MCL**	M	1.5	0.6
		SD	5.2	4.2
		Range	-8.7-15.3	-6.7-14.7
	BNL	M	1.3	0.5
		SD	5.0	6.3
		Range	-9.3-14.0	-16.0-20.0
	ANL	M	0.3	0.1
		SD	4.5	4.9
		Range	-12.7-9.3	-15.3-12.7
		CR	8.8	9.7



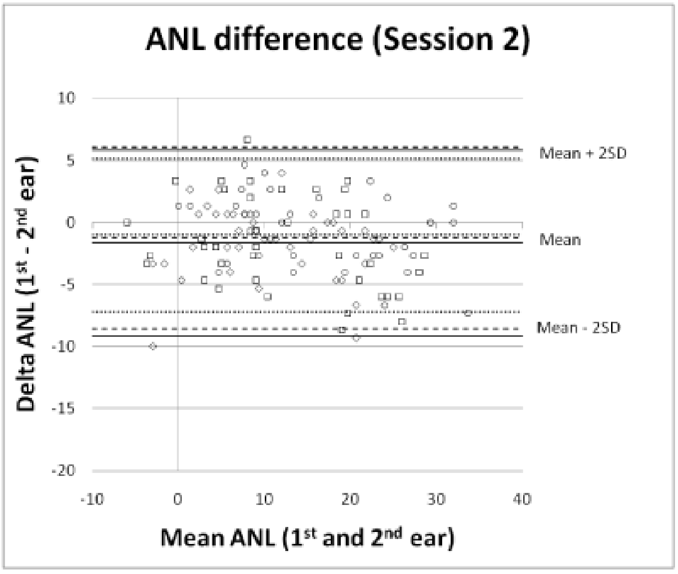
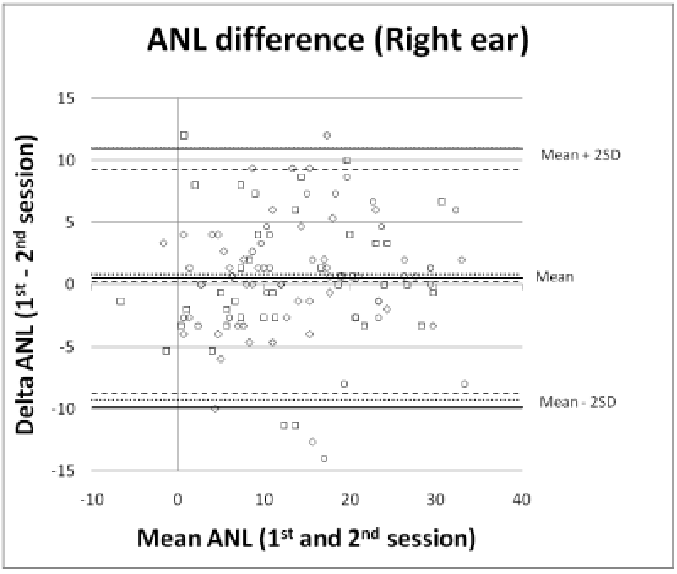


Figure 1.
Bland-Altman plots showing the differences between ANL measurements at the first and the second ear separated in time by approximately 15-30 minutes. The differences are plotted as a function of the mean of the two ANL measurements for three conditions. Data are shown separately for the two sessions. Data from condition 1 are shown as squares (mean, + 2 SD, -2 SD are shown as continuous lines), data from condition 2 are shown as circles (mean, + 2 SD, -2 SD are shown as dotted lines) and data from condition 3 are shown as diamonds (mean, + 2 SD, -2 SD are shown as dashed lines).



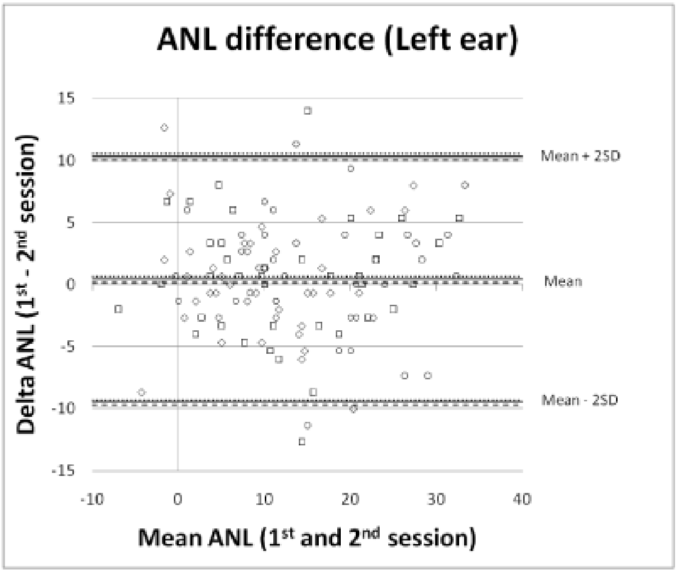


Figure 2.
Same as figures 1, but showing the differences between ANL measurements at the first and the last session as a function of the mean ANL obtained at the two sessions.