The North American Listening in Spatialized Noise - Sentences Test (NA LISN-S): Normative data and test-retest reliability studies for adolescents and young adults

David Brown#+
Sharon Cameron*†
Jeffrey Martin~
Charlene Watson‡
Harvey Dillon*

*National Acoustic Laboratories
†Macquarie University
Cincinnati Children's Hospital Medical Center
+University of Cincinnati
~University of Texas at Dallas
‡Community Audiology Services, Calgary Health Region

Corresponding Author:

David Brown

The test described in this article is distributed under licensed by Phonak Communications AG. Financial returns from the sale of this product will benefit Dr. Cameron and the organizations involved in this study. This outcome has in no way influenced the research reported in this article.

ABSTRACT

Background: The Listening in Spatialized Noise – Sentences test (LISN-S; Cameron & Dillon, 2009) was originally developed to assess auditory stream segregation skills in children aged 6 to 11 years with suspected central auditory processing disorder. The LISN-S creates a three-dimensional auditory environment under headphones. A simple repetition-response protocol is used to assess a listener's speech reception threshold (SRT) for target sentences presented in competing speech maskers. Performance is measured as the improvement in SRT in dB gained when either pitch, spatial, or both pitch and spatial cues are incorporated in the maskers. A North American-accented version of the LiSN-S (NA LiSN-S) is available for use in the United States of America and Canada.

Purpose: To develop normative data for adolescents and adults on the NA LiSN-S, to compare these data with those of children aged 6 to 11 years as documented in Cameron et al. (2009), and to consolidate the child, adolescent and adult normative and retest data to allow the software to be used with a wider population.

Research Design: In a descriptive design, normative data and test-retest reliability data were collected.

Study Sample: One hundred and twenty normally-hearing participants took part in the normative data study (67 adolescents aged 12 years, 1 month to 17 years, 10 months and 53 adults aged 19 years, 10 months to 30 years, 30 months). Forty-nine participants returned between one-and-four months after the initial assessment for retesting. Participants were recruited from sites in Cincinnati, Dallas and Calgary.

Results: When combined with data collected from children aged 6 to 11 years, a trend of improved performance as a function of increasing age was found across performance measures. ANOVA revealed a significant effect of age on performance ($p < 0.00001$ for all measures). Planned contrasts revealed that there were no significant differences between adults and children aged 13 years and older on the low cue SRT; 14 years and older on talker and spatial advantage; 15 years and older on total advantage; and 16 years and older on the high cue SRT. Mean test-retest differences on the various NA LISN-S performance measures for the combined child, adult and adolescent data ranged from 0.05 dB to 0.5 dB. Paired comparisons revealed test-retest differences were not significant on any measure of the NA LISN-S (p ranging from 0.100 to 0.822) except low cue SRT ($p = 0.003$). Test-retest differences across measures did not differ as a function of age ($p = 0.307$ to 0.973). Test and retest scores were significantly correlated for all NA LiSN-S measures (*r* ranging from 0.3 to 0.6 and p ranging from 0.006 to < 0.000001).

Conclusions: The ability to use either spatial or talker cues in isolation becomes adult-like by about 14 years of age, whereas the ability to combine spatial and talker cues does not fully mature until closer to adulthood. By consolidating child, adolescent, adult normative and retest data the NA-LISN-S can now been utilized to assess auditory processing skills in a wider population.

Key words: auditory stream segregation; central auditory processing disorder; spatial hearing loss.

Abbreviations: CAPD = central auditory processing disorder; HpTF = headphone transfer function; HRTF = head related transfer function; LISN-S = Listening in Spatialized Noise – Sentences test; NA LISN-S = North American Listening in Spatialized Noise – Sentences test; SD = standard deviation; SNR = signal-to-noise ratio; SHL; spatial hearing loss; SRT = speech reception threshold.

The following paper describes two studies conducted to collect normative and test-retest reliability data from adolescents and adults for the North American-accented version of the Listening in Spatialized Noise – Sentences test (LISN-S; Cameron and Dillon, 2009). The LISN-S was developed to assess the various skills involved in auditory stream segregation in children with suspected central auditory processing disorder (CAPD). Auditory stream segregation is the process by which a listener is able to differentiate the various auditory signals which arrive simultaneously at the ears and form meaningful representations of the incoming acoustic signals (Sussman et al, 1999). Auditory cues such as the perceived spatial location of sounds, or the pitch of speakers' voices, help this process of segregating the total stream of sound (Bregman, 1990).

A description of the LiSN-S detailing the development and evaluation of this audiological assessment tool is available in the literature (Cameron and Dillon, 2007a and 2007b; Cameron & Dillon 2008a and 2008b; Cameron et al. 2009). A brief overview follows.

The LISN-S is presented using a personal computer. Output levels are directly controlled by the software via an external USB soundcard. A three-dimensional auditory environment under headphones is created by pre-synthesizing the speech stimuli with head-related transfer functions (HRTFs). A simple repetition-response protocol is used to assess a listener's speech reception threshold (SRT) for target sentences presented in competing speech maskers (children's stories). Using HRTFs, the targets are perceived as coming from directly in front of the listener (0° azimuth) whereas the maskers vary according to either their perceived spatial location (0° vs. + and -90° azimuth), the vocal identity of the speaker/s of the stories (same as, or different to, the speaker of the target sentences), or both. This test configuration results in four listening conditions: same voice at 0º (or low cue SRT); same voice at ±90º; different voices at 0º; and different voices at ±90º (or high cue SRT).

Performance on the LISN-S is evaluated on the low and high cue SRT, as well as on three "advantage" measures. These advantage measures represent the benefit in dB gained when either talker (pitch), spatial, or both talker and spatial cues are incorporated in the maskers, compared to the baseline (low cue SRT) condition where no cues are present in the maskers (see Figure 1). The use of relative measures of performance (i.e., difference scores) serves to minimize the influence of higher order language, learning and communication skills on test performance. For example, as such skills affect both the SRT when the distracters are presented at 0°, and the SRT when they are spatially separated at ±90°, these skills will have minimal effect on the difference in dB between the SRTs in these two conditions. Thus, the differences that inevitably exist between individuals in such functions can be accounted for, allowing for clearer evaluation of their abilities to use spatial and voice cues to aid speech understanding.

The LISN-S has shown to be sensitive to auditory streaming deficits in children with suspected CAPD whose primary difficulties in the classroom stem from poor listening behaviour (Cameron and Dillon, 2008a). For these children, significant differences on LISN-S occurred only in the conditions where the physical location of the maskers was different to that of the target speaker (high cue SRT, $p = 0.001$; spatial advantage, $p < 0.0001$, and total advantage, $p < 0.0001$), suggesting that the children specifically exhibited a disorder that we could term spatial hearing loss (SHL). Normative and test-retest reliability data on both the Australian and North American LiSN-S currently exists for children aged 6 years, 0 months to 11 years, 11 months.

Picard and Bradley (2001) reviewed 11 studies on the deleterious effects of noise on spoken word recognition by normally-hearing children and young adults. The sample mean recognition scores in noise were regressed on age to explore possible developmental changes with age. There was a clear trend for spoken word recognition in noise to develop with increasing age for normally-hearing participants, irrespective of response format.

In order to investigate auditory stream segregation skills, and in particular spatial hearing loss, in older children and adults on the LiSN-S it is therefore necessary to collect additional normative and retest reliability data. This paper describes the collection and analysis of such data from participants aged 12 to 30 years of age. Comparison of results with that of children aged 6 to 11 years (as documented in Cameron et al. 2009) is also described. NA LiSN-S cut-off scores and critical difference scores (used to determine whether a genuine improvement in performance on the LiSN-S has been achieved following some period of management or remediation) have been recalculated with the combined child, adolescent and adult data and documented in this paper.

EXPERIMENT 1 - NORMATIVE DATA STUDY

Approval to conduct the normative data study was obtained from the Institutional Review Boards of the Cincinnati Children’s Hospital Medical Center, the University of Texas at Dallas, and the University of Calgary.

Participants

Data were collected from a total of 120 participants, comprising 67 adolescents (aged 12 years, 1 month to 17 years, 10 months) and from 53 adults (aged 19 years, 10 months to 30 years, 3 months). There were 56 males and 64 females. Full participant details are provided in Table 1. Participants were recruited from friends and family of staff at the Cincinnati Children's Hospital and Calgary Health Region. Adolescent participants recruited by the University of Texas at Dallas were from local schools. The adult participants were recruited from the university and received course credit for their participation. The participants were included in the study if they had North American English as a first language, no history of hearing disorders, and no reported learning or attention disorders. On the day of testing all participants had pure tone thresholds of ≤ 15 dB HL at 500 to 4000 Hz, and ≤ 20 dB HL at 250 and 8000 Hz, as well as normal Type A tympanograms.

Design and Procedure

Testing was carried out in an acoustically treated room suitable for testing hearing thresholds at the various facilities between 9 am and 6 pm.

Materials

The LISN-S stimuli were administered using a PC and Sennheiser HD215 headphones. The headphones were connected directly to the headphone socket of the PC via a Miglia Harmony Express USB soundcard. The sensitivity of the soundcard was automatically set to a pre-determined level by the LISN-S software in order to achieve pre-designated signal levels.

This alleviated the need for daily calibration. At this pre-set level, the combined distracters at 0˚ had a long-term rms level of 55 SPL as measured in a Brüel and Kjær type 4153 artificial ear.

Target sentences were initially presented at a level of 62 dB SPL. Competing children's stories, looped during playback, were presented at a constant level of 55 dB SPL (for the combined level of the two competing talkers). The target and competing signals were presented simultaneously to both ears. The listener's task was to repeat the words heard in each target sentence. A 1000 Hz 200 ms tone burst was presented before each sentence to alert the listener that a sentence would be presented. A silent gap of 500 ms separated the tone burst from the onset of the sentence. The tone burst was presented at a constant level of 55 dB SPL. The instructions provided to each participant are detailed in Cameron et al. (2009). Up to 30 sentences were presented in each of the four conditions of distracter location and voice: same voice at 0º (SV0º), same voice at ±90º (SV±90º), different voices at 0º (DV0º) and different voices at ±90º (DV±90º).

The SNR was adjusted adaptively in each condition by varying the target level to determine each participant's SRT. The SNR was decreased by 2 dB if a listener scored more than 50 percent of words correct, and increased by 2 dB if he or she scored less than 50 percent of words correct. The SNR was not adjusted if a response of exactly 50 percent correct was recorded (for example, 3 out of 6 words correctly identified). All words in each sentence were scored individually; including the definite article "the" and the indefinite articles "a" and "an". A minimum of five sentences were provided as practice, however practice continued until one upward reversal in performance (i.e. the sentence score dropped below 50 percent of words correct) was recorded. Testing ceased in a particular condition when the listener had either (a) completed the entire 30 sentences in any one condition; or (b) completed the practice sentences plus a minimum of a further 17 scored sentences, and their standard error, calculated automatically in real time over the scored sentences, was less than 1 dB.

The presentation order of the LISN-S conditions was counterbalanced among participants using a Latin-square protocol to enable analysis of the effect of practice on performance.

RESULTS

Where appropriate in the following analyses, data from the participants aged 12 to adult from the present study were combined with data from children aged 6 to 11 from Cameron et al (2009) in order to ascertain the effects of maturation on NA LiSN-S performance, and to calculate revised NA LiSN-S cut-off scores. Analyses were performed with Statistica 7.1.

Effect of Data Collection Site

The mean SRT and inter-participant standard deviations (SD) for the NA LISN-S SRT and advantage measures for each collection site for the current study are presented in Table 2. Separate ANOVAs were performed to determine the effect of collection site (Ohio, Calgary and Texas) on each of the performance measures. As the five measures were derived from the four basic LISN-S conditions (SV0º, SV±90º, DV0º and DV±90º), the alpha level of 0.05 was multiplied by 4/5 to give an adjusted level of 0.04 to avoid inflating the Type I error rate.

There was no effect of collection site on the high cue SRT, $F(2, 117) = 2.045$, $p = 0.134$; talker advantage, $F(2, 117) = 0.411$, $p = 0.664$; spatial advantage, $F(2, 117) = 0.886$, $p = 0.415$; or total advantage, $F(2, 117) = 0.822$, $p = 0.442$. There was a significant difference on the low cue SRT, $F(2, 117) = 3.780$, $p = 0.026$. Post hoc tests using Tuckey's HSD revealed that Ohio and Calgary were just significantly different from each other ($p = 0.035$) on the low cue SRT measure. The mean difference in dB between Ohio and Calgary on the low cue SRT was only 0.5 dB (-1.8 dB in Ohio vs. -1.3 dB in Calgary) and the SD for each centre was minimal (0.9 dB for Ohio and Texas and 1.0 dB for Calgary). Based on the small difference between centres on the low cue SRT, and the fact that there were no significant differences between centres on any other LiSN-S performance measure, the adult and adolescent data from the three participating centres were combined for all further analyses, as was the case with the child data in Cameron et al (2009).

Effect of Age on NA LISN-S Performance Measures

The mean SRT and advantage measures in the normative data study are illustrated in Figure 2 (a) to (e). There was a trend of decreasing SRT and increasing advantage, as age (6 to adult) increased, across measures. Separate ANOVAs were performed to determine the effect of age on the performance measures. As for previous analyses, the alpha level of 0.05 was multiplied by 4/5 to give an adjusted level of 0.04 to avoid inflating the Type I error rate.

For the low cue SRT there was a significant main effect of age, $F(12, 179) = 7.640$, $p < 0.000001$. Planned comparisons revealed significant differences between adults and children 12 years and younger. There was also a significant main effect of age for the high cue SRT, $F(12, 179) = 14.527$, $p < 0.00001$. Planned comparisons revealed significant differences between adults and children 15 years and younger.

There was a significant main effect of age for the talker advantage measure, $F(12, 179) = 10.353$, $p <0.00001$. A significant main effect of age was also found for the spatial advantage measure, $F(12, 179) = 9.724$, $p <0.000001$. For both the talker and spatial advantage measures planned comparisons revealed significant differences between adults and children 13 years and younger. Finally, there was a significant main effect of age for the total advantage measure, $F(12, 179) = 7.4418$, $p < 0.000001$. Planned comparisons revealed significant differences between adults and children 14 years and younger.

Gender Effects

An analysis was conducted in order to investigate gender effects across age groups. Mean scores and SDs for the 94 females and 98 males on the various LISN-S SRT and advantage measures are provided in Table 3, along with the results of ANOVAs which were performed with each measure as the dependant variable, a fixed factor of gender, and age as a covariate. There was no significant effect of gender for any LISN-S measure (p ranging from 0.179 to 0.834).

Practice Effects

As presentation order was counterbalanced in the adolescent and adult normative data study, the mean SRTs were compared for the four basic LISN-S conditions (SV0º, SV±90º, DV0º and DV±90º) as a function of presentation order (first, second, third or fourth) in order to determine whether practice improved performance. Age groups (12 to adult) were combined to provide sufficient numbers in each condition and task combination to calculate meaningful inferential statistics. Table 4 shows the mean thresholds in dB for each LISN-S condition as a function of presentation order.

One way ANOVAs revealed no significant differences in mean SRTs as a factor of presentation order for either the SV0° condition, $F(3,116) = 0.639$, $p = 0.591$; the SV±*90°*, $F(3,116) = 1.865$, $p = 0.139$; or the DV 0° condition, $F(3,116) = 0.712$, $p = 0.547$. There was, however, a significant difference for the DV ±90° condition, $F(3,116) = 4.677$, $p = 0.004$. Post hoc tests using Tukey's HSD revealed that, at -13.9 dB, the first presentation resulted in a significantly higher SRT than the second, third and fourth presentations at -15.3 dB, -15.3 dB and -15.5 dB respectively ($p < 0.05$). Adjustments to cut-off scores to account for the effect of change of presentation order on performance when the test is presented clinically across age groups are detailed in the following sections.

Regression Analysis and LISN-S Cut-Off Scores

As a strong trend of improved performance with increasing age was found for the various LISN-S SRT and advantage measures, it was determined that cut-off scores, calculated as two SDs below the mean, would need to be adjusted for age for each performance measure. These cut-off scores represent the level below which performance on the LISN-S is considered to be outside normal limits.

A regression analysis utilising a two-piece fit was conducted with SRT or advantage for each measure as the independent variable and age (ranging from 6.21 to 30.3 years) as the dependent variable. The cut-off scores were adjusted for age using the formulas:

1. Low and High Cue SRT: max (intercept + B * age, c) + (2 * SDs of the residuals from the age-corrected trend lines) + correction factor (if applicable)

2. Talker, Spatial and Total Advantage: min (intercept + B * age, c) – (2 * SDs of the residuals from the age-corrected trend lines) – correction factor (if applicable)

All regression data are presented in Table 5. Figure 3 (a) to (e) provides scatterplots of the regression analysis showing the individual data points. The correction factors utilized to account for the effect of presentation order on performance are described below.

Effect of Change of Presentation Order on Performance

In the study by Cameron et al (2009) with children aged 6 to 11 years, the first presentation of the LISN-S DV0º condition resulted in a SRT that was significantly higher than for the third presentation. It was therefore recommended that in clinical practice the presentation order of the LiSN-S conditions should be (1) DV±90º; (2) SV±90º; (3) DV0º; and (4) SV0º. This configuration represents a gradient from “easy” to “difficult”, and controls for the practice effects that were demonstrated for the DV0º condition in the Cameron et al (2009) study.

However, as the presentation order of conditions in the Cameron et al (2009) study was counterbalanced, only a quarter of the participants received the LISN-S in the recommended order described above. In order to determine the effect of change of presentation order on performance to that recommended for clinical use, a regression analysis was conducted with presentation order (1, 2, 3 or 4) as the dependent variable and SRT in each LISN-S condition as the independent variable. A significant effect of order was found for the DV±90º condition. The adjustment to the normative data needed to test in the recommended order was calculated as the number of steps away from the order midpoint (2.5), multiplied by the B-value (dB/step). A 0.9 dB adjustment was required for the DV±90º (high cue SRT) condition. As the SNR of the DV±90˚ condition is also utilized in the calculation of the total advantage score, a correction factor of -0.9 dB was implemented for the total advantage measure (Cameron et al, 2009).

An examination of practice effects in the current adolescent and adult normative data study found a significant difference between the first and subsequent presentations of the DV±90º condition – but not the DV0 º condition. For consistency, the presentation order recommended in Cameron et al (2009) will be retained for the combined child, adolescent and adult normative data, and as noted in the preceding section a correction factor will be applied to the revised cut-off scores to account for the effect of change of presentation order on NA LiSN-S performance.

In order to determine the correction faction for the current study, a regression analysis was again conducted as for the Cameron et al (2009) study. The adjustment to the adolescent and adult normative data needed to test in the recommended order was 0.7 dB for the DV±90˚ condition and 0 dB for the SV±90º, DV0º and SV0º conditions. The mean correction factor for the combined child, adolescent and adult data was thus +0.8 dB for the DV±90˚ condition (high cue SRT) and -0.8 dB for the total advantage measure. The adjustments are noted in Table 5.

EXPERIMENT 2 – TEST-RETEST RELIABILITY STUDY

Approval to conduct the test-retest reliability study was obtained from the Institutional Review Boards of the Cincinnati Children's Hospital Medical Center, the University of Texas at Dallas, and the University of Calgary.

Participants

Data were collected by Cincinnati Children's Hospital, Calgary Health Region and the University of Texas at Dallas. Participants were 49 of the 120 participants who had taken part in the adolescent and adult normative data study who agreed to also take part in the test-retest reliability study. Participant details are provided in Table 6. Participants were tested between 9 am and 6 pm.

Design and Procedure

The materials and procedures used in the test-retest reliability study were as for the normative data study. The four LISN-S conditions were presented to each participant in the same order that they were presented during the normative data study. Retesting on the LISN-S in the adolescent and adult retest reliability study was carried out between 0 months, 24 days to 4 months, 7 days following the initial testing (mean 2 months, 14 days, median 2 months, 3 days).

RESULTS

In order to examine trends in performance as a function of age, and to facilitate recalculation of the NA LiSN-S critical difference scores across age groups, all of the following analyses were undertaken with the child retest reliability study data from Cameron et al (2009) and the current adolescent and adult retest reliability study data combined. Analyses were performed with Statistica 7.1.

Test-Retest Paired Comparisons

The mean scores and SDs for the various LISN-S conditions - and the advantage measures derived from the various conditions - at test and retest, are provided in Table 7. Differences in scores in dB between test and retest, as well as the t- and p-values of paired-samples t-tests are also provided. All differences were in the direction representing an improvement in performance except for the Spatial Advantage measure where the paired difference was only 0.05 dB. Averaged across all participants, the maximum improvement in performance on retest was 0.8

dB on the DV0˚ condition. Minimum change was 0.1dB on the spatial advantage and total advantage conditions. There was a significant differences in performance between test and retest on the low-cue SRT condition (paired difference 0.5 dB, $p = 0.003$) and the DV0˚ condition (paired difference 0.8 dB, $p = 0.003$). There were no significant differences between test and retest score on any other LiSN-S measure.

Effect of Age on LISN-S Performance Measures

Figure 4 (a) to (e) depicts the mean test and retest scores for each of the LISN-S SRT and advantage measures, as a function of age (6 to adult). A repeated measures analysis of variance (ANOVA) was performed for each NA LISN-S SRT and advantage measure, with age as a between-participants factor, to determine whether test-retest differences differed with age. An alpha level of 0.05 was used for all comparisons. There was no significant interaction of test session and age for the low-cue SRT, $F(12,72) = 0.36$, $p = 0.974$; the high-cue SRT, $F(12,72) = 0.44$, $p = 0.942$; talker advantage measure, $F(12,72) = 1.19$, $p = 0.307$; spatial advantage, $F(12,72) = 0.89$, $p = 0.632$; or total advantage, $F(12,72) = 0.49$, $p = 0.915$.

Test-Retest Correlation Analysis

A Pearson product-moment correlation analysis between test and retest scores was performed for each of the NA LISN-S SRT and advantage measures. All correlations were significant. For the low-cue SRT, $r = 0.3$, $p = 0.006$; high-cue SRT, $r = 0.6$, $p < 0.00001$; talker advantage, $r = 0.5$, $p < 0.000001$; spatial advantage, $r = 0.5$, p $p < 0.000001$; and for total advantage, $r = 0.5$, $p < 0.0001$. Scatter plots in Figure 5 (a) to (e) show the correlation of test versus retest scores for each of the LISN-S SRT and advantage measures.

Test-Retest Correction Factors

Table 8 displays the calculations of the one-sided critical differences required to determine whether an individual has improved on the LISN-S following remediation or compensation, taking the mean test-retest differences into account. That is, for the increase in an individual advantage measure score to be significant at the 95 percent confidence interval level, the score must increase by 1.64 x SD of the test-retest difference plus the mean test-retest difference. Similarly, for a significant improvement on an SRT measure score to have occurred, the magnitude of the decrease must be greater than 1.64 x SD of the test-retest difference plus the mean test-retest difference. Across measures, the standard deviation of paired test-retest differences were similar for adolescents, adults and children. Critical difference measures, including the correction factor, ranged from 3.0 dB on the low cue SRT to 4.4 dB on the high cue SRT.

DISCUSSION

The North American LISN-S (NA LISN-S) adolescent and adult normative data and test-retest reliability studies were carried out in line with the design and procedures employed in the NA LISN-S child studies (Cameron et al, 2009).

The normative data study by Cameron et al (2009) showed performance on the NA LiSN-S improved as a function of age for children aged 6 to 11 years. In the present study performance continued to improve with increasing age for individuals aged 12 years to 30 years. ANOVA revealed a significant effect of age with p values less than 0.00001 across all LiSN-S measures. Planned contrasts revealed that there were no significant differences between adults and children aged 13 years and older on the low cue SRT; 14 years and older on talker and spatial advantage; 15 years and older on total advantage; and 16 years and older on the high cue SRT.

As the presentation order on NA LiSN-S conditions was counterbalanced in both the present study and in Cameron et al (2009), the effect of practice on the NA LISN-S was examined by measuring the effect of position within the four subtests on the performance of each subtest. For the DV ±90º condition there was a significant effect of test order between the first and subsequent positions (p=0.004). There was no significant effect of practice on any other NA LISN-S condition (p ranging from 0.591 to 0.139).

However, clinically, the LiSN-S is presented in a specific order (DV±90º; SV±90º; DV0º then SV0º) as recommended in Cameron et al (2009). Only a quarter of the participants in both the present study and in Cameron et al (2009) study received the LiSN-S in this recommended order due to counterbalancing of the presentation order of conditions during the collection of the normative data. A regression analysis was therefore conducted to determine the effect of changing the presentation order to that recommended by Cameron et al (2009). There was a significant effect of presentation order for the DV±90º condition (p = 0.002). An adjustment to the normative data needed to account for the delivery of the LISN-S conditions in the recommended order was calculated as the number of steps away from the order midpoint (2.5), multiplied by the B-value (dB/step). Across all age groups, an adjustment of 0.8 dB was needed to compensate for presenting the DV±90º (high cue SRT) condition first. This adjustment was also applied to the total advantage measure which is consequently affected as it is calculated as the difference between the SV0 º and DV±90º conditions. These correction factors were applied to the NA LiSN-S cut-off scores.

The NA LiSN–S cut-off scores represent the level below which performance on a particular measure is considered to be outside normal limits of plus or minus two standard deviations. A regression equation was utilized to calculate the cut-off scores as a function of age. A two-piece regression line was fitted to the data from individuals aged 6.21 to 30.3 years for each NA LiSN-S SRT and advantage measure. The correction factor (to account for change of presentation order from that used in the normative data study to that recommended clinically) was added to the formula for the high cue SRT and total advantage cut-off scores.

The NA LISN-S test-retest reliability study revealed that mean test-retest differences on the various NA LISN-S conditions and performance measures for the combined child, adult and adolescent data ranged from 0.05 dB on the spatial advantage measure to 0.8 dB on the different voices 0° condition. Across all age groups (6 to 30 years), paired comparisons revealed test-retest differences were not significant on any measure of the NA LISN-S (p ranging from 0.100 to 0.822) except the low cue SRT ($p = 0.003$), for which the mean difference was 0.5 dB. Further, mean test-retest differences across measures did not differ as a function of age ($p = 0.307$ to 0.973) and test and retest scores were significantly correlated for all NA LiSN-S measures (*r* ranging from 0.3 to 0.6 and p ranging from 0.006 to < 0.000001).

As per Cameron et al (2009) the test-retest data was utilized to develop one-sided critical difference scores applicable across age groups (6 years to adult). These scores can be used to determine whether an individual has genuinely improved on the LISN-S following a period of remediation or compensation with an assistive listening device. Critical differences ranged from 3.0 dB on the low cue SRT to 4.4 dB on the high cue SRT.

CONCLUSION

In previous studies (Cameron et al, 2009; Cameron & Dillon, 2007a, 2007b, 2008), the LISN-S was reported to be a fast and efficient assessment tool with potential to be used clinically to evaluate auditory stream segregation skills – particularly spatial hearing loss - in children aged 6 to 11 years with suspected CAPD in Australia and North America. The present study has described the development of normative and retest reliability data for older children and young adults from the United States of America and Canada. An expected trend of improved performance as a function of increasing age was found for all measures. The ability to use spatial or talker cues in isolation becomes adult-like by about 14 years of age, whereas the ability to combine spatial and talker cues does not fully mature until closer to adulthood. The NA LiSN-S cut-off scores, which determine the level below which performance on the NA LISN-S is considered to be outside normal limits, were adjusted for age accordingly. The calculation of one-sided critical difference scores, which take into account mean practice effects and day-to-day fluctuation in performance, makes the NA-LISN-S a potentially valuable tool for monitoring performance over time and the effects of maturation, remediation, or compensation such as an assistive listening device.

ACKNOWLEDGEMENTS

The authors would like to thank Cari Olsen and Jennifer Wright from the University of Cincinnati who collected the data for the normative data study and test-retest reliability study for that facility.

REFERENCES

Bregman AS. (1990) Auditory scene analysis. Cambridge: The MIT Press.

Cameron S, Brown D., Keith R, Martin J, Watson, C, Dillon, H. (2009). Development of the North American Listening in Spatialized Noise - Sentences Test (NA LISN-S): Sentence equivalence, normative data and test-retest reliability studies. *J Am Acad Audiol*, 20(2):128-146.

Cameron S, Dillon H. (2009) *Listening in Spatialized Noise – Sentences test (LISN-S)* (Version 1.014) [Computer software]. Murten, Switzerland: Phonak Communications AG.

Cameron S, Dillon H. (2008a). The Listening in Spatialized Noise – Sentences Test: Comparison to prototype LISN test and results from children with either a suspected (central) auditory processing disorder of a confirmed language disorder. *J Am Acad Audiol*, *19*(5):377-391.

Cameron S. & Dillon, H. (2008b). Spatial hearing deficits as a major cause of auditory processing disorders: Diagnosis with the LISN-S and management options. In R. Seewald & J. Bamford, eds. *A Sound Foundation Through Early Amplification 2007. Proceedings of the Fourth International Conference*: Phonak AG, Switzerland, 235-241.

Cameron S, Dillon, H. (2007a) Development of the Listening in Spatialized Noise - Sentences Test (LISN-S). *Ear Hear* 28(2):196-211.

Cameron S, Dillon H. (2007b) The Listening in Spatialized Noise - Sentences Test (LISN-S): Test-retest reliability study. *Int J Audiol 46*:145-153.

Picard M, Bradley J. (2001). Revisiting speech interference in classrooms. *Audiology, 40*:221-244.

Sussman E, Ritter W, Vaughan HG. (1999) An investigation of the auditory streaming effect using event-related brain potentials. *Psychophysiolog,* 36:22–34.

Figure 1. The LISN-S SRT and advantage measures.

Figure 2. Dot plots of the normative data for the 120 adolescents and adults in the current study and the 72 children in the children in the Cameron et al (2009) study on the LiSN-S SRT (a and b) and advantage measures (c-e). Error bars represent the 95 percent confidence intervals from the mean.

Figure 3. Scatterplots of the normative data for the 120 adolescents and adults in the current study and the 72 children in the children in the Cameron et al (2009) study on the LiSN-S SRT (a and b) and advantage measures (c-e). Non-linear regression lines were fitted for the low and high cue SRT measures as max (intercept + B-value * age, c); and for the spatial, talker and total advantage measures as min (intercept + B-value * age, c).

Figure 4. Dot plots depicting mean test and retest scores for each of the NA LISN-S SRT (a and b) and advantage measures (c to e), as a function of age. Circular filled symbols connected by solid lines represent the test scores. Square unfilled symbols connected by dashed lines represent the retest scores. Error bars represent the 95% confidence intervals from the mean. The figure incorporates data from the 36 children from Cameron et al (2009) and the 49 adolescents and adults from the current study.

Figure 5. Scatter plots depicting correlation of test vs. retest scores for each of the LISN-S SRT (a and b) and advantage measures (c to e). Lines show the 95% confidence interval of the regression line. The figure incorporates data from the 36 children from Cameron et al (2009) and the 49 adolescents and adults from the current study.