The Development of the Listening in Spatialized Noise – Universal Test (LiSN-U) and Preliminary Evaluation in English-Speaking Listeners

Sharon Cameron[a, b], Kiri Mealings[a], Nicky Chong-White[a], Taegan Young[a], and Harvey Dillon[a, b, c]

[a]*National Acoustic Laboratories, Sydney, Australia*

[b]*Macquarie University, Sydney, Australia*

[c]*University of Manchester, Manchester, United Kingdom*

Keywords: auditory processing disorder, spatial processing disorder, diagnostic tools, speech perception

Acronyms

ANOVA – Analysis of variance

dB – Decibel

CAPD – Central auditory processing disorder

CV – Consonant-vowel

DV0 – LiSN-S Different Voices 0° condition

DV90 – LiSN-S Different Voices ±90° condition (also known as the high cue condition)

ESL – English as a Second Language

HRTF – Head-related transfer function

LiSN-S – Listening in Spatialised Noise – Sentences Test

LiSN-U – Listening in Spatialised Noise – Universal Test

RMS – Root mean square

SNR – Signal-to-noise ratio

SPD – Spatial processing disorder

SPL – Sound pressure level

SRT – Speech reception threshold

SV90 – LiSN-S Same Voice ±90° condition

SV0 – LiSN-S Same Voice 0° condition (also known as the low cue condition)

aSE – Adjusted standard error

Corresponding Author:

Dr Kiri Mealings

National Acoustic Laboratories

Level 5 Australian Hearing Hub

16 University Avenue

Macquarie University, NSW, 2109, Australia

Email: kiri.mealings@nal.gov.au

The Development of the Listening in Spatialized Noise – Universal Test (LiSN-U) and Preliminary Evaluation in English-Speaking Listeners

Objective: To create a language independent version of the Listening in Spatialized Noise – Sentences test (LiSN-S) and evaluate it in an English-speaking population.

Design: Test development and normative data collection. LiSN-Universal (LiSN-U) targets consisted of CVCV pseudo-words (e.g. /mupa/). Two looped distracter tracks consisted of CVCVCVCV pseudo-words. The listener's task was to repeat back the target pseudo-words. Stimuli were presented over headphones using an iPad. Speech reception thresholds were measured adaptively. In the co-located condition all stimuli came from directly in front. In the spatially-separated condition the distracters emanated from +90° and -90° azimuth. Perceived location was manipulated using head-related transfer functions. Spatial advantage was calculated as the difference in dB between the co-located and spatially-separated conditions.

Study Sample: Stimulus intelligibility data were collected from 20 adults. Normative data were collected from native English speakers (23 adults and 127 children).

Results: Children's spatially-separated, co-located, and spatial advantage results improved significantly with age. Spatial advantage was 4-6 dB larger in the LiSN-U than LiSN-S depending on age group.

Conclusions: Whereas additional research in non-native English populations is required, the LiSN-U appears to be an effective tool for measuring spatial processing ability.

Introduction

Central auditory processing disorders (CAPD) is an umbrella term for a variety of impairments related to poor perceptual processing of auditory information in the central auditory nervous system despite a person having normal hearing thresholds (American Speech-Language-Hearing Association, 2005). Spatial processing disorder (SPD) refers to a specific type of CAPD that results from an inability to use time and intensity differences of auditory signals between the ears coming from various locations to segregate a target signal from competing signals (Cameron & Dillon, 2007a, 2008, 2011, Cameron, Dillon, & Newall, 2005, 2006b; Cameron, Glyde, & Dillon, 2012; Glyde, Cameron, Dillon, Hickson, & Seeto, 2013; Glyde, Hickson, Cameron, & Dillon, 2011). Children with SPD need a higher signal-to-noise ratio (SNR) to achieve the same speech reception thresholds (SRTs) compared to children without the disorder. SPD is diagnosed using the Listening in Spatialized Noise – Sentences test (LiSN-S) (Cameron & Dillon, 2009). Research over the last 15 years has shown that a significant number of children with reported listening difficulties relative to their peers have a SPD (Cameron & Dillon, 2008, 2011; Cameron et al., 2012; Cameron, Glyde, Dillon, King, & Gillies, 2015). For example, of the 666 children who received a CAPD assessment by the Australian Hearing CAPD service over an 18 month period, 130 (19.5% of the sample) were diagnosed with SPD (Cameron et al., 2015).

The LiSN-S measures the SNR at which a listener can segregate a target sentence from simultaneously presented competing speech. The LiSN-S speech stimuli are convolved with head-related transfer functions to create a three-dimensional auditory environment over headphones. To effectively segregate the speech signals, the listener needs to be able to utilise the inter-aural time and level spatial cues of the speech between ears. The LiSN-S is described in detail in Cameron and Dillon (2007a).

In summary, the test has four conditions which are compared to determine whether or not a child has SPD (Cameron & Dillon, 2007a, 2007b, 2008). The four conditions are Different Voices ±90° condition (DV90, also known as the high cue condition), Same Voice ±90° condition (SV90), Different Voices 0° condition (DV0), and Same Voice 0° condition (SV0, also known as the low cue condition). The target voice comes from the front (0° azimuth) in all conditions, but the distractors differ in either being the same or different from the target, and whether they are presented at 0° azimuth or ±90° azimuth. The LiSN-S software calculates the SRTs and z scores for each of these conditions, and also calculates the talker advantage, spatial advantage, and total advantage comparative scores which represent the benefit gained when either vocal properties (such as pitch), spatial, or both vocal and spatial cues are incorporated in the competing speech, compared to a baseline condition where neither cues are present in the competing speech. The use of relative measures of performance reduces the influence of higher-order communication skills on performance. Thus, the differences that inevitably exist between individuals can be controlled for, allowing for clearer evaluation of their abilities to use spatial cues to aid speech understanding.

Research with the LiSN-S has shown that children and adults with normal hearing experience a spatial advantage when the competing speech comes simultaneously from both sides (±90° azimuth) compared with when both target sentences and competing speech all come from directly in front of the listener (0° azimuth). The size of this advantage (on average 12.5 dB across participants, SD = 2.1 dB) occurs when the target and competing speech are spoken by the same female speaker, i.e. in the absence of any pitch differences between voices (Cameron & Dillon, 2011). Spatial advantage improves with age, with six year olds experiencing an average advantage of 10 dB compared to 14 dB for young adults. However, children with spatial processing disorder need a significantly greater SNR in order to achieve the same SRT as normal-hearing controls. A listener with SPD would typically perform well below the average for their age (i.e. z < -2) on the DV90 condition and the SV90 condition, but perform within the average range on the DV0 and SV0 conditions. This pattern of results would give them a poor spatial advantage score but a typical talker advantage and total advantage score.

Various studies have linked SPD with severity and duration of chronic otitis media (Graydon, Rance, Dowell, & Van Dun, 2017; Tomlin & Rance, 2014). The proportion of children with SPD rises to 7% in Indigenous Australian children (Cameron, Dillon, Glyde, Kanthan, & Kania, 2014), which is a population that experiences pervasive chronic otitis media from an early age. Fortunately, SPD can be reversed with deficit-specific auditory training using the Sound Storm iOS program (Cameron & Dillon, 2011; Cameron et al., 2012, 2015).

The LiSN-S speech material is in Australian-English (Cameron & Dillon, 2007a) and American-English (Cameron, Brown, & Keith, 2009) accents and each version uses words familiar to speakers of those accents. However, it may be problematic for clients with other English accents, and for clients who do not have English as their first language as speech perception abilities can be influenced by the accent of the speaker relative to the listener. Flowerdew (1992) provides an overview of several studies that demonstrate comprehension difficulty for both native and non-native listeners when listening to unfamiliar native and/or non-native accents. There is also strong evidence that speech in a familiar accent is easier to understand than if it is in an unfamiliar accent (Major, Fitzmaurice, Bunta, & Balasubramanian, 2002). It is actually whether the accent is familiar to the listener rather than whether the accent is similar to the listener's that aids listening comprehension (Major et al., 2002; Tauroza &

Luk, 1997). The addition of noise can further impair the intelligibility of accented speech (Munro, 1998).

A person's native language also affects their speech perception abilities. A listener's English proficiency has been shown to significantly affect their ability to recognise speech in noise. Kilman, Zekveld, Hällgren, and Rönnberg (2014) reported that SRTs are better when the listener is listening to their native language in noise compared to their non-native language. Tabri, Chacra, and Pring (2011) found that while monolingual speakers and bilingual speakers perform equally well on speech perception tasks in quiet conditions, bilingual speakers have greater speech perception difficulties in noise at poor SNRs (i.e. +5 dB SNR and 0 dB SNR). Rogers, Lister, Febo, Besing, and Abrams (2006) also found significantly poorer word recognition scores for bilingual compared to monolingual listeners in noise and noise with reverberation, but not in quiet. Krizman, Bradlow, Lam, and Kraus (2016) found that Spanish–English bilinguals were worse than English monolingual adolescents at perceiving sentences and words, but better than English monolinguals when perceiving tones in noise. Skoe and Karayanidi (2018) found that bilingual Chinese speakers were less able to take advantage of contextual cues than monolingual native American-English speakers for a test of sentences in noise. Furthermore, Cameron, Barker, and Newall (2003) demonstrated that having English as a second language (ESL) can affect performance on English-language CAPD tests, as the language barrier can result in some children with ESL performing at levels associated with CAPD. However, this is due to the language barrier rather than true CAPD. LiSN-S is, of course, completely unsuitable for people who do not speak English at all. A recent study by Schafer et al. (2018) examined the presence of spatial cues in adults and children who spoke American English or Mandarin Chinese as a native language with the LiSN-S. The authors found that performance was better for the native English speakers in nearly every test condition.

Therefore, there is a need for speech perception tests able to detect SPD to be delivered in a person’s native language and accent. There have been requests to develop the LiSN-S in different languages and dialects. However, establishing the LiSN-S in a variety of languages involves recreation, recording, equalisation, and validation of the speech materials, as well as collection of normative test/retest data and clinical validation.

Due to the issues raised above, a universal version of the LiSN-S was developed, that uses made-up words comprised of consonants and vowels that are common to many languages, so people from any language background can be tested for SPD. Hyman's (2007) consonantal and vocalic universal rules were followed. Consonant-vowel (CV) segments were chosen to make up the two-syllable speech segments as most languages have CV syllables due to the sonority cycle principal which states that the preferred syllable structure has a sonority profile that “rises maximally towards the peak and falls minimally towards the end” (Clements, 1990, pg 301; Hyman, 2007).

The aim of this study was to create a new test for detecting SPD, called the Listening in Spatialized Noise – Universal Test (LiSN-U) that is language independent. This study evaluates the LiSN-U in an English-speaking population. Future studies will evaluate its use with people from other language backgrounds with the aim for it to be used to diagnose people from any language background with SPD, rather than needing to create new versions of the LiSN-S for each language. The LiSN-U has the potential to take the success of the LiSN-S and make this technology available to children and adults around the world who have difficulty hearing in background noise relative to their peers due to SPD. It will also allow for better diagnoses of Aboriginal and Torres Strait Islanders who are more susceptible to SPD due to prolonged and recurrent otitis media, but who often have English as a second, or even third, language, particularly if they are from remote locations. This paper describes the development of the LiSN-U test, the stimulus intelligibility study, and the normative data collection results with English-speaking adults and children. Test-retest reliability of the LiSN-U and a direct comparison of LiSN-S and LiSN-U performance in a group of children aged six and seven years will be described in future papers.

Method: Development of the LiSN-U Test Materials

CV Stimuli

The consonants and vowels selected for the LiSN-U were chosen because of the large number of languages in which they are used ("Phonemic Inventories and Cultural and Linguistic Information Across Languages", "The Speech Accent Archive"). The consonants finally adopted comprised /p, b, t, d, k, g, m, n, s, h/. The consonants /f, v, z, l/ were also originally selected but were excluded after the intelligibility study. The vowels consisted of the close front vowel /i/, the open front vowel /a/, and the close back vowel /u/.

CV Recording, Editing, and Convolution

Each of the original 14 consonants were paired with each of the three vowels (e.g. /bi/; /ba/; /bu) to create 42 CV pairs. The CVs were recorded by a female speaker in a general Australian accent. General Australian is the stereotypical variety of Australian English used by the majority of Australians and it dominates the accents found in contemporary Australian-made films and television programs. Recording took place in a chamber, anechoic above 50 Hz. The stimuli were recorded on a personal computer using Adobe Audition version C5.6, a RME Babyface Pro USB audio interface (Haimhausen, Germany) and a Rode NT1 cardioid condenser microphone (Silverwater, Australia) with a pop guard. The recordings were edited using Adobe Audition C5.6. Each CV was saved as an individual audio file. Audio was recorded at 32 bit at a sample rate of 48 kHz. Each CV was then level normalized to have a root mean square (RMS) level of -20 dB re: digital full scale. The average length of the CVs was 420 ms (range 400 – 439 ms). The 42 individual target CVs were then convolved with HRTFs at 0° azimuth (modified version of HRTFs recorded by Cameron, Dillon, and Newall, 2006a) and level normalized to a Total RMS level of -30 dB (in the left and right channels). We wanted to allow data collection using any high-quality headphone, so no additional processing was performed to accommodate the transfer function of a particular headphone. The 42 target CV's as described above were utilized in the Stimulus Intelligibility Study.

Target Stimuli Generation

Two CV audio files were concatenated to make the CVCV tokens (i.e. with no gaps between the stimuli), for example /muba/. Forty-two CV_1CV_2 target tokens were randomly generated so that there were equal occurrences of consonants and vowels and each of the 14 consonants occurred three times. CV_1 and CV_2 were not allowed to be the same, although the same consonant could appear in both CVs, as could the same vowel.

Distracter Track Generation

Two distracter tracks were generated using the 42 non-convolved CV audio files. The distracter tracks were made up of 42 multi-syllable *pseudo-words*, i.e. CV_1-CV_2-CV_3-CV_4. For example /ba/ /di/ /mu/ /sa/. Syllable combinations were randomly selected, with the restriction that each CV occurs in each position of the pseudo-word only once. CV_1 and CV_2 were concatenated and extended to 880 ms by inserting silence after CV_2 so that there is a gap between CV_2 and CV_3. For example, /badi musa/. CV_3 and CV_4 were concatenated and extended to 1080 ms (880 ms + 200 ms), by inserting silence after CV_4 so that there is a gap between CV_4 and CV_1 of the following pseudo-word. The start time of one of the distracter tracks was delayed so that the gaps in distracter_1 do not occur at the same time as the gaps in distractor_2. Both distractor tracks (of 42 *pseudo-words* each) were level normalized to -30 dB RMS re full digital scale, and then spatialized to either 0° azimuth or +90° and -90° azimuth.

As the target and distractor signals are spoken by the same talker in both conditions, we refer to the two LiSN-U conditions just as spatially-separated and collocated.

Experiment 1: Stimulus Intelligibility Study

Participants

The participants were 20 adults aged 21 yr, 5 m to 38 yr, 7 m (mean 33 yr, 1 m). There were 9 females and 11 males. The most proficient language for each participant was English x 10, Dutch x 1, French x 3; Croatian x 1; Mandarin x 1; Spanish x 2; Mauritian (Creole) x 1; and Hebrew x 1. Ethics approval was obtained from the Australian Hearing Human Research Ethics Committee. Participants signed a consent form before commencing the study. Testing took place in an audiometric test booth at the National Acoustic Laboratories.

Pure tone audiometric screening was performed using an Interacoustics AC4 clinical audiometer (Middelfart, Denmark) with circumaural Sennheiser HDA 200 audiometric headphones (Hanover, Germany). All participants had normal hearing, defined as equal to, or better than, 20 dB HL at all octave frequencies from 500 to 8000 Hz measured bilaterally.

Materials

The LiSN-U software was developed in Xcode Version 8.2.1 using Swift 3.1 programming language (Apple Inc, 2017, California, USA). The LiSN-U test materials were presented using an Apple Air 2 iPad running iOS 10.2.1 (California, USA) and Sennheiser HD 200 Pro circumaural headphones (Hanover, Germany). The sensitivity of the iPad sound card was automatically set by the LiSN-U software. At this pre-set level the combined distracters at 0° have a long-term root mean square (RMS) level of 65 dB sound pressure level (SPL) as measured in a Brüel and Kjær Head and Torso Simulator type 4128-C-001 (Nærum, Denmark).

Target CVCV Presentation

A 200 millisecond (ms) 1kHz warning tone was presented at +5 dB SNR (70 dB SPL), followed by a 300 ms silent gap. The target CVCV token was then presented followed by a 200 ms gap and then repeated, for example: warning tone – 300 ms – /bima/ – 200 ms – /bima/.

Scoring

One point was scored for each C correct in any order, and one point was scored for each V correct in any order, with a maximum of four points per trial. It was decided to score a phoneme correct regardless of order as the pseudo-words were repeated in each trial with only a short 200 ms gap between items. As such the listener may hear the last two phonemes of the pseudo-word followed by the first two phonemes and respond, for example, with /guti/ for /tigu-tigu/. The LiSN-S does not require the words to be repeated in the correct order. The disadvantage with this method is that the chance score will be higher than if order was required. However, because we are targeting the 75% point, and because there are three vowels and nine consonants, the chance score will still be much lower than the point on the psychometric function that we are targeting, so will have only a minor effect on the threshold measured.

Practice and Test Conditions

The practice and test phases were completed in two conditions. The first condition was the spatially-separated condition, where the target tokens are presented at 0° azimuth, distracter track 1 is presented at +90° azimuth, and distracter track 2 is presented at -90° azimuth. The second condition was the co-located condition, where the target tokens are presented at 0° azimuth, and distracter track 1 and 2 are both presented at 0° azimuth.

Procedure

Distracter Track Playback Levels

When adding two non-correlated/incoherent waveforms of equal level, the total RMS level increases by 3 dB. As such the distracter$_1$ and distracter$_2$ tracks were each played simultaneously during the practice and test phases at a level 3 dB lower than the desired distractor level. Consequently, the total RMS level (i.e. of distractor$_1$ + distractor$_2$) equaled the total average RMS level of the target CVCVs at a 0 dB SNR.

Practice Phase

The distracter track level was set at a constant level of 65 dB SPL. Target CVCV tokens were randomly presented (as described in the *Target CVCV Presentation* section) starting at an SNR of +11 dB (76 dB SPL) and adjusted adaptively. Six CVCV trials were presented for the practice phase. Practice was given prior to commencement of both the spatially-separated and co-located conditions.

Test Phase and Target Step Sizes

The distracter track level was set at a constant level of 65 dB SPL. The initial target presentation level was at +11 dB SNR (76 dB SPL) and adjusted adaptively. If no phonemes were correct the target increased by 3 dB. If one phoneme was correct the target increased by 2 dB. If two phonemes were correct the target increased by 1 dB. If three phonemes were correct there was no change in the target level. If all four phonemes were correct the target decreased by 4 dB *before* the first upward reversal in the spatially-separated condition, and decreased by 2 dB *after* the first upward reversal in the spatially-separated condition. In the co-located condition, if four phonemes were correct, the target decreased by 2 dB irrespective of the first upward reversal.

Calculation of Speech Reception Threshold

The speech reception threshold for each participant was calculated as:

- SRT = mean(target levels after the first reversal, and including the next target level that would be presented) - (background noise level, i.e. 65 dB SPL)

For the Stimulus Intelligibility Study, the SRT was calculated for each of the 42 target CV tokens presented.

Results: Stimulus Intelligibility Study

A psychometric function was fitted to the combined scored data from all participants in both the spatially-separated and the co-located conditions for each consonant and vowel. Three consonants (/f/, /l/ and /v/) were removed as their percent correct scores and psychometric functions indicated that they were much harder than the other consonants, and /z/ was removed because its psychometric function indicated it was too easy.

Target Stimuli Adjustment – Post-Stimulus Intelligibility Study

Based on the results of the Stimulus Intelligibility Study, as reflected in the psychometric functions, the level of the remaining 30 CV audio files at 0° azimuth were adjusted based on the difference between the SNR at the 50% point and the average across all phonemes of these SNRs. The adjustment for each CV was equal to the weighted average of the adjustment needed for the consonant and the adjustment needed for the vowel. The averaging weights were based on the ratio of the average error rates for the ten included consonants (37.8%) and for the three vowels (15.9%). Adjustments ranged from +1.5 dB for /hi/ and /ha/ to -1.5 dB for /tu/. The adjustments were made so that the RMS level at 0° azimuth, averaged across all 30 CVs, remained unchanged at -30 dB re full digital scale in both the left and right channels.

Distracter Track Generation - Post-Intelligibility Equalization

Two new distracter tracks were generated using the non-convolved CV audio files (at RMS -30 dB) of the 30 CV target tokens as determined by the Stimulus Intelligibility Study, and were convolved with HRTFs at 0° and ±90° to produce the two distractor files.

Experiment 2: Normative Data Collection

Participants

The participants were 23 adults aged 19 yr 5 m to 56 yr 5 m (mean 30 yr 1 m) and 127 children aged 5 yr 0 m to 12 yr 0 m (mean 8 yr 8 m). For the adults there were 13 females and 10 males. For the children there were 63 females and 64 males. All participants were native English speakers. The children included were those whose parents reported that they did not have an attention, language, or learning problem. An additional 13 children participated in the study but six were excluded due to inattention or difficulty completing the task, one was a clear outlier based on their spatial advantage score, and the other six children were removed from the data analysis as their adjusted standard errors (aSE; based on the variability within the adaptive track; Cameron and Dillon, 2007b) for the spatially-separated and/or spatial advantage were large and were considered outliers, based on the distribution of aSE outliers on the scatterplots and corroborated by an examination of the relevant adaptive tracks. In the majority of cases the aSE outliers were the direct result of a premature commencement of the measurement phase in the spatially-separated condition for reasons other than reaching threshold level, e.g. needing more practice or a momentary lapse of attention. This has since been changed so that the practice phase is longer to avoid this issue (see Discussion section).

The child participants completed the testing in a quiet room at their primary school. Sound levels in the school testing rooms were measured between 45-50 dBA using a Q1362 digital sound level meter (Dick Smith Electronics, Australia). Pure tone audiometric screening was conducted using an Interacoustics Audio Traveler A222 portable audiometer (Middelfart, Denmark). Parental consent was given for all children who took part in the study. For the consenting adults, testing took place in an audiometric test booth at the National Acoustic Laboratories. Audiometric screening was performed using an Interacoustics AC4 clinical audiometer (Middelfart, Denmark) with circumaural Sennheiser HDA 200 audiometric headphones (Hanover, Germany). All participants had normal hearing, defined as equal to, or better than, 20 dB HL at all octave frequencies from 500 to 8000 Hz measured bilaterally.

Materials

The LiSN-U software was redeveloped in Xcode Version 8.3.2 using Swift 3.1 programming language. Copyright 1999-2017 Apple Inc (California, USA).

Procedure

The procedure was as per the Stimulus Intelligibility Study with the following exceptions. For the familiarisation phase, the target CVCV tokens were presented in quiet at a fixed level of 75 dB SPL. Five CVCV tokens were created containing all ten consonants paired with a vowel (so each vowel occurred at least three times). Five consonants occurred in CV_1 position, and the remaining five consonants occurred in CV_2 position. Each CVCV token was preceded by a 200 ms 1 kHz warning tone followed by a 300 ms silent gap. A minimum of five trials (CVCV stimuli) were presented. If an error occurred in any CVCV token, feedback was provided and the token was repeated until correctly identified.

In the practice phase, the initial target presentation level was changed to +10 dB SNR (75 dB SPL). Five trials were presented, rather than six. In the co-located condition, if all four consonants and vowels were correctly identified the SNR reduced by 2 dB from the first target token presented, rather than by 4 dB, as the typical threshold was much closer to the initial test level than was the case for the spatially-separated condition. In the test phase, the maximum number of trials (CVCV tokens) presented was 30, with equal numbers of consonants and vowels. Additionally, a stopping criterion was implemented if the aSE of the mean was less than 1.0 dB and a minimum of 17 measurement trials (i.e. trials occurring after the first upward reversal) has been completed.

All participants were encouraged to give an answer to each stimulus even if it was a guess. If they did not hear any part of the CVCV pseudo-word they could say pass.

Data Analysis

Data analysis was completed using Statistica version 13. Analyses included an examination of the effects of age and gender on the participants' spatially-separated SRT, co-located SRT, and spatial advantage. A regression and residuals analysis was also conducted. This enabled z score formulas to be derived and added to the app so that future testing would classify whether a listener's performance is around the mean value or above/below the mean on the test conditions when compared to their peers.

Results: Normative Data Collection

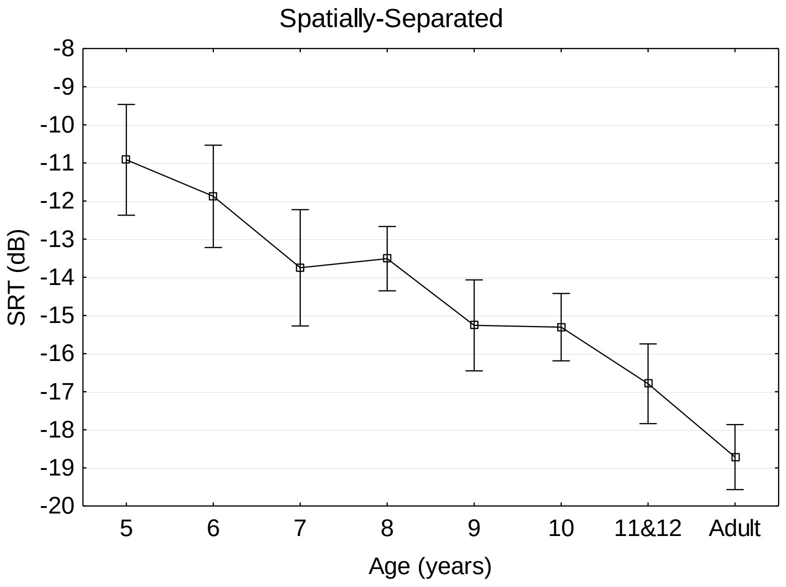
Examination of Age and Gender

Separate ANOVAs were conducted for each of three dependent measures: spatially-separated SRT, co-located SRT, and spatial advantage, all quantified in dB. In all three models, age was a continuous predictor, and gender was a categorical predictor. There was a significant effect of age on each dependent measure, but no effect of gender on any of the measures. The ANOVA results are shown in Table 1. The mean performance by age group is shown graphically in Figure 1. There were 13 5-year-olds, 19 6-year-olds, 19 7-year-olds, 18 8-year-olds, 18 9-year-olds, 22 10-year-olds, and 17 11-year-olds, one 12-year-old, and 23 adults. Post hoc Tukey HSD tests were used to determine significant differences between age groups. For the spatially-separated condition, the mean SRT for the 5-year-olds was significantly different to participants aged seven and older. The mean for the 6-year-olds was significantly different to those aged nine and older. The means for the 7-year-olds and the 8-year-olds were significantly different to those aged ten and older. The means for the 9-year-olds and the 10-year-olds were significantly different to the adults. The mean for the 11-12-year-olds was not significantly different to the adults. For the co-located condition, the mean SRT for the 5-year-olds was significantly different to participants aged nine and older. The mean for the 6-year-olds was significantly different to those aged ten and older. The means for the 7-year-olds and the 8-year-olds were significantly different to those aged eleven and older. The means for the 9-year-olds, the 10-year-olds, and the 11-12-year-olds were not significantly different to the adults. For the spatial advantage, the means for the 5-year-olds and the 6-year-olds were significantly different to participants aged nine and older. The means for the 7-year-olds was significantly different to the adults. The means for the 8-year-olds was significantly different to the participants aged 11 and older. The means for the 9-year-olds and the 10-year-olds were significantly different to the adults. The mean for the 11-12-year-olds were not significantly different to the adults.

Table 1: Age and gender ANOVA results for spatially-separated, co-located, and spatial advantage measures.

	SS	*df*	*MS*	*F*	*p*
Intercept	7946	1	7946	1050	<0.0005

Spatially-separated	Age	505	1	505	67	<0.0005
	Gender	9	1	9	1	0.29
	Error	1112	147	8		
Co-located	Intercept	463	1	463	267	<0.0005
	Age	35	1	35	20	<0.0005
	Gender	1	1	1	0.6	0.43
	Error	255	147	2		
Spatial advantage	Intercept	12229	1	12229	2103	<0.0005
	Age	277	1	277	48	<0.0005
	Gender	3	1	3	0.6	0.45
	Error	855	147	6		



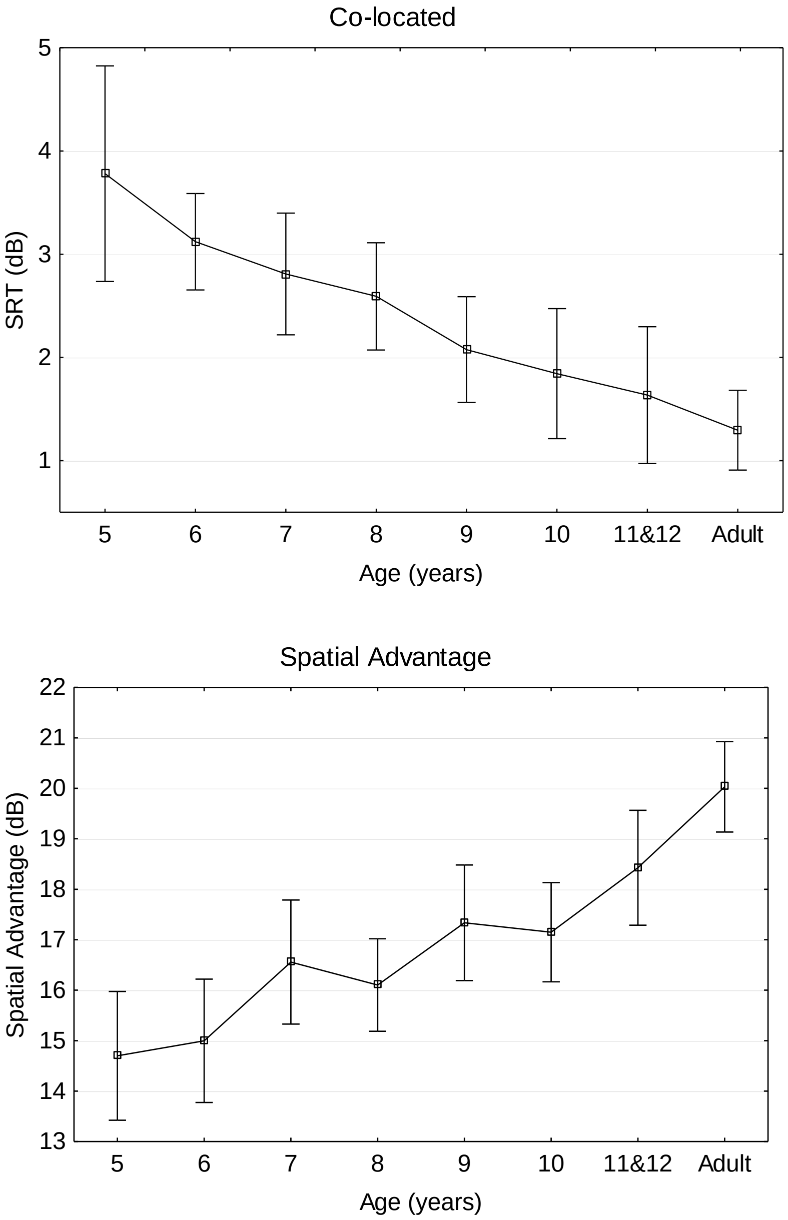


Figure 1: Graph showing mean SRTs on spatially-separated and co-located conditions, and spatial advantage measure by age group. Bars represent 95% confidence intervals.

Regression Analysis, Residuals Analysis, and Calculation of z Scores

A regression and residuals analysis was conducted so formulas for z score calculations could be derived. As documented in Tables 2 and 3, the variances shown in Figure 1 were larger for the children than for the adults. The variances were significantly different for children and adults for the spatially-separated condition ($F(126, 22) = 2.41$, $p < 0.01$) and the co-located condition ($F(126, 22) = 2.44$, $p < 0.01$). As such, residuals were calculated separately for children and adults. Figure 2a), b), and c) show the regression of the spatially-separated, co-located, and spatial advantage scores in dB versus age for the children.

Table 2 shows the analysis of the residuals for the children using the formula:

- Measure residual = Measure - (a + b*Age)

The standard deviation (SD) of the residuals shows how much spread there is around the regression line of SRT versus age. The SD for each measure is shown in Table 2 for the children and in Table 3 for the adults. The distance of any individual data point from the regression line (for the children) or from the mean (for the adults) is divided by the appropriate standard deviation to produce the z score.

The adult and child data were fitted with a two-segment linear fit as follows:

- $z_{\text{spatially-separated, co-located}}$ = -(score in dB - Max (a + b*Age, c)) / SD of residuals
- $z_{\text{spatial advantage}}$ = (score in dB - Min (a + b*Age, c)) / SD of residuals

where c is the mean score for adult performance. The combined fit enables this equation to be applied to people aged from approximately 5 to 50 years.

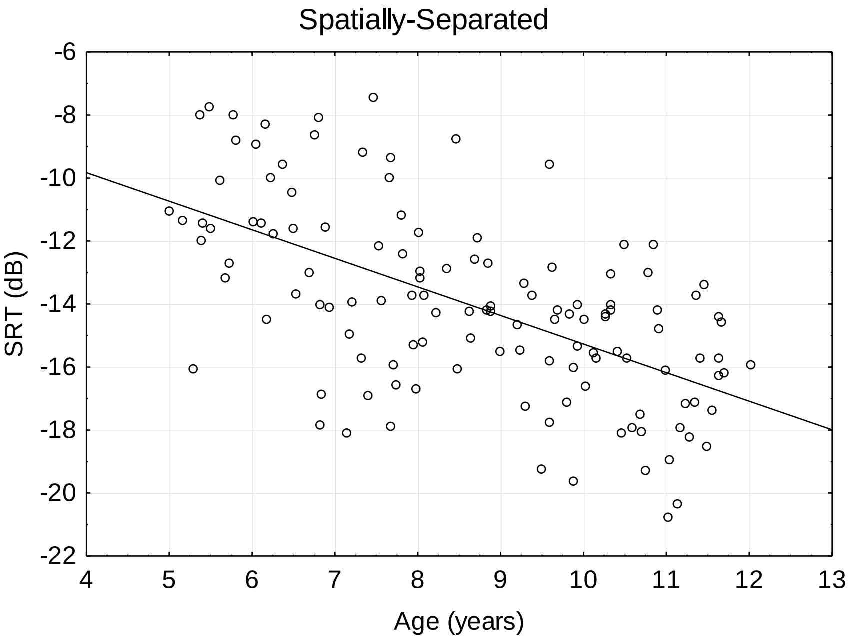
Table 2: Analysis of residual for the spatially-separated, co-located, and spatial advantage data for 127 children. The correlation coefficient r shows the correlation of each measure with age.

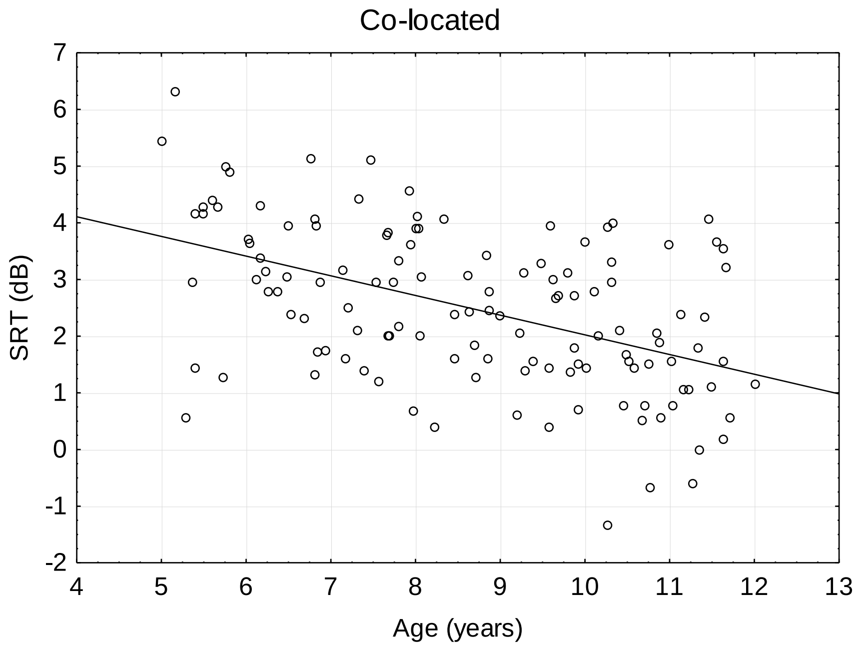
Measure	*r*	*a*	*b*	*Var*	*SD*	*SD* of residual
Spatially-separated	-0.596	-6.2	-0.907	9.394	3.06	2.46

Co-located	-0.486	5.5	-0.347	1.934	1.39	1.22
Spatial advantage	0.433	11.7	0.560	6.955	2.64	2.38

Table 3: Means and standard deviations for spatially-separated, co-located, and spatial advantage data for 23 adults.

Measure	*Mean*	*SD*
Spatially-separated	-18.7	1.97
Co-located	1.3	0.89
Spatial advantage	20.0	2.07





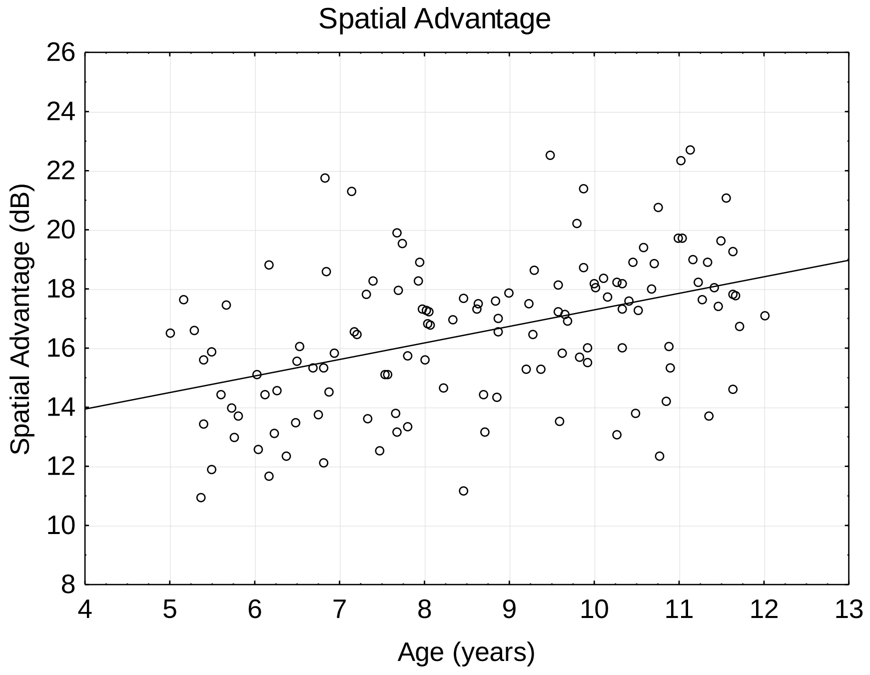


Figure 2: Scatterplots of the a) spatially-separated, b) co-located, and c) spatial advantage scores in dB as a function of age for the 127 children in the present study.

Comparison of LiSN-S and LiSN-U Spatial Advantage by Age

In order to compare the LiSN-U and LiSN-S spatial advantage performance a piece-wise linear regression was fitted to the LiSN-U spatial advantage scores (in dB) from the 23 adults and 127 children in the present study, as well as previously published LiSN-S normative data from 203 children and adults (Cameron & Dillon, 2007a; reported in Cameron, Glyde, & Dillon, 2011) as shown in Figure 3. Because age is plotted on a logarithmic axis the sloping lines are curved. Spatial advantage is approximately 4-6 dB larger in the LiSN-U than in LiSN-S depending on the age group.

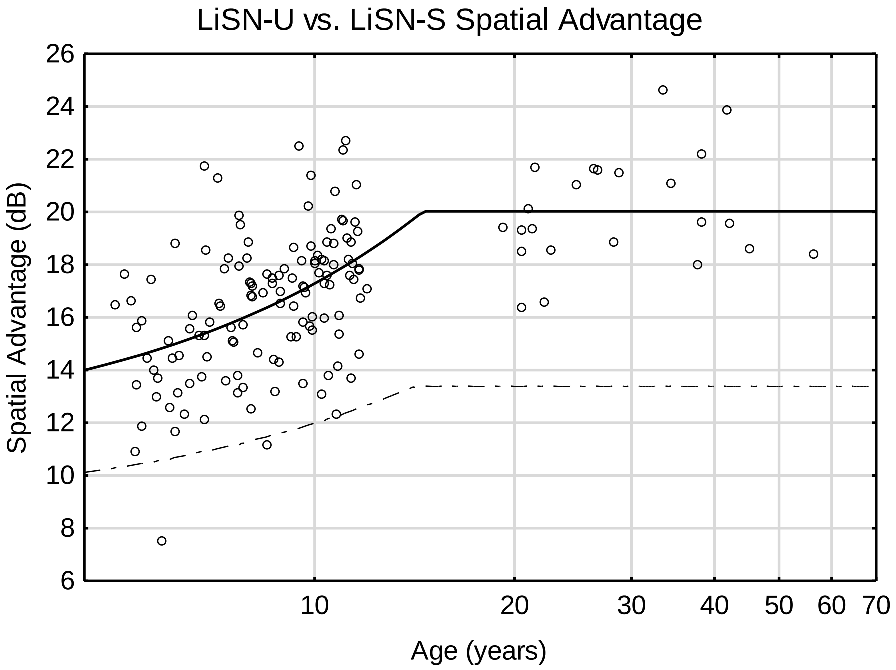


Figure 3: LiSN-U and LiSN-S spatial advantage comparison for both adults and children. The dots represent the LiSN-U scores. The solid line represents the piece-wise linear regression for the LiSN-U (n = 150 from the present study). The dashed line represents the piece-wise linear regression from previously published normative data for the LiSN-S (n = 202).

Discussion

The aim of this study was to develop a new language-independent version of the LiSN-S test of spatial processing, and collect normative data in an English-speaking population. The advantage of this new test is that it could be used in the future to diagnose SPD in people from language backgrounds other than English, including Indigenous populations where the higher prevalence of otitis media can often lead to SPD, without needing to translate the LiSN-S into many different languages. The phonemes chosen for the stimuli were those common to many different languages, and these were refined in the stimulus intelligibility study to give 10 consonants and 3 vowels that make up the CVCV pseudo-words for the distractor tracks and target stimuli.

As found in the LiSN-S, significant improvements in the SRTs as the participant's age increased were found for the LiSN-U spatially separated and co-located conditions, as well as the spatial advantage measure. Spatial advantage was approximately 4 to 6 dB larger in the LiSN-U than in LiSN-S depending on the age group. Possibly this is because there are fewer differences between target and distractor in the LiSN-U than in LiSN-S. If so, this might make the LiSN-U an even more sensitive detector of SPD than the LiSN-S. As noted in the normative data study method section, the results of some children were excluded due to aSE outliers resulting from variability in the adaptive track. This variability was a direct result of a premature commencement of the measurement phase in the spatially-separated condition. The length of the practice phase was increased in this condition from 4 to 6 presentations to alleviate this problem for clinical use.

The inter-participant variances were greater for children compared to adults, significantly so ($p < 0.01$) for the spatially-separated condition and the co-located condition, but not significantly so for spatial advantage. The greater inter-participant variance for children in the two baseline conditions may be because a range of abilities, all of which potentially develop at different rates across children, will affect the baseline scores. The spatial advantage measure, being a difference measure, is much less affected by all the abilities that affect both baseline conditions. Alternative reasons for greater inter-participant variance for the children (in any of the scores) are that the children were tested at school rather than in a sound-attenuating booth like the adults and that the same HRTFs were used for the children as the adults, rather than adjusting them for head size. In creating the LiSN-U, it was also hoped that the test would be able to be taken by slightly younger children than the LiSN-S as it has fewer language demands. The LiSN-S is recommended for children aged six and above due to variability in performance for younger ages. Data on the LiSN-U were collected with five-year-olds, however there was higher variability on co-located condition for this age group compared to other ages (but not in the spatially-separated or spatial advantage measure), see Figure 2. Therefore, care should be taken when using the LiSN-U with this age group to make sure the children understand the test and can stay attentive throughout.

For this study, phonemes were considered correct if they were repeated back in any order. As mentioned, it was decided to score a phoneme correct regardless of order as the pseudo-words were repeated in each trial with only a short 200 ms gap between items. As such the listener may hear the last two phonemes of the pseudo-word followed by the first two phonemes. The LiSN-S does not require the words to be repeated in the correct order so the same approach was taken for this test. Future research, however, could assess the differences between this scoring method and only scoring phonemes correct if they are in the right order.

Conclusions and limitations

The present research has shown that the LiSN-U is an easy-to-use test for measuring spatial processing in native English-speaking children and adults. Although the test has been designed to be applicable to talkers of almost any language, the appropriateness of the normative data to children whose first language is other than Australian English should be ascertained before the z-scores calculated by the test could be relied upon. It is possible that idiosyncrasies of the talker accent used to create the stimuli could impact performance for non-Australian-English speakers. If so, we would expect that the spatial advantage measure, being a difference score, would be less affected than the base measures. Future research is also needed to assess the ability of the LiSN-U to detect SPD in a clinical population.

Diagnosis of SPD in countries and communities where English is not the first language will allow those affected by SPD to potentially access government or private sector assistance in managing the disorder, thereby potentially reducing the risk of academic and social issues arising from an inability to adequately access auditory input in environments such as the classroom. The test takes 10-12 minutes to administer and results are readily available on completion.

Acknowledgements

The authors would like to thank the NSW Department of Education and the primary schools who took part in this research. The participation of the children and their families are also appreciated. This research is funded by the Australian Government through the Department of Health. Harvey Dillon acknowledges the support of Macquarie University, the NIHR Manchester Biomedical Research Centre, and the Australian Department of Health.

Professional meeting details

Cameron, S., Mealings, K., Chong-White, N., & Dillon, H. (2018, May). Development of the Listening in Spatialized Noise – Universal test (LiSN-U). A language-independent test for spatial processing disorder. *Paper presented at* the *Audiology Australia National Conference 2018, Sydney, Australia.*

Disclosure statement. The authors report no conflicts of interest.

References

American Speech-Language-Hearing Association. (2005). *(Central) auditory processing disorders [Technical statement]*. Retrieved from www.asha.org/policy

Cameron, S., Barker, R., & Newall, P. (2003). The effect of linguistic background on the Macquarie Pediatric Speech Intelligibility Test. *The Australian and New Zealand Journal of Audiology*, *25*(2), 95–98.

Cameron, S., Brown, D., & Keith, R. (2009). Development of the North American Listening in Spatialized Noise-Sentences test (NA LiSN-S): Sentence equivalence, normative data, and test-retest reliability. *Journal of the American Academy of AUdiology*, *20*(2), 128–146. doi:10.3766/jaaa.20.2.6

Cameron, S., & Dillon, H. (2007a). Development of the Listening in Spatialized Noise-Sentences Test (LISN-S). *Ear and Hearing*, *28*(2), 196–211. doi:10.1097/AUD.0b013e318031267f

Cameron, S., & Dillon, H. (2007b). The listening in spatialized noise-sentences test (LISN-S): Test-retest reliability study. *International Journal of Audiology, 46*, 145–153. doi:10.1080/14992020601164170

Cameron, S., & Dillon, H. (2008). The Listening in Spatialized Noise–Sentences Test (LISN-S): Comparison to the prototype LISN and results from children with either a suspected (central) auditory processing disorder or a confirmed language disorder. *Journal of the American Academy of Audiology*, *19*(5), 377–391.

Cameron, S., & Dillon, H. (2009). Listening in Spatialized Noise – Sentences test (LiSN-S) (Version 2.4). Murten, Switzerland: Phonak Communications AG.

Cameron, S., & Dillon, H. (2011). Development and evaluation of the LiSN & Learn Auditory Training Software for deficit-specific remediation of binaural processing deficits in children: Preliminary findings. *Journal of the American Academy of Audiology*, *22*(10), 678–696. doi:10.3766/jaaa.22.10.6

Cameron, S., Dillon, H., Glyde, H., Kanthan, S., & Kania, A. (2014). Prevalence and remediation of spatial processing disorder (SPD) in Indigenous children in regional Australia. *International Journal of Audiology*, (November 2013), 326–335. doi:10.3109/14992027.2013.871388

Cameron, S., Dillon, H., & Newall, P. (2005). Three case studies of children with suspected auditory processing disorder. *Australian and New Zealand Journal of Audiology*, *27*(2), 97–112.

Cameron, S., Dillon, H., & Newall, P. (2006a). Development and evaluation of the Listening in Spatialized Noise test. *Ear and Hearing*, *27*(1), 30–42.

Cameron, S., Dillon, H., & Newall, P. (2006b). The Listening in Spatialized Noise test: An auditory processing disorder study. *Journal of the American Academy of Audiology*, *17*(5), 304–318.

Cameron, S., Glyde, H., & Dillon, H. (2011). Listening in Spatialized Noise-Sentences Test (LiSN-S): Normative and retest reliability data for adolescents and adults up to 60 years of age. *Journal of the American Academy of Audiology*, *22*(10), 697–709. doi:10.3766/jaaa.22.10.7

Cameron, S., Glyde, H., & Dillon, H. (2012). Efficacy of the LiSN & Learn auditory training software: Randomized blinded controlled study. *Audiology Research, 2*(e15), 86–93. doi:10.4081/audiores.2012.e15

Cameron, S., Glyde, H., Dillon, H., King, A., & Gillies, K. (2015). Results from a national central auditory processing disorder service: A real-world assessment of diagnostic practices and remediation for central auditory processing disorder. *Seminars in Hearing*, *36*(4), 216–236. doi:10.1055/s-0035-1564457

Clements, G. N. (1990). The role of the sonority cycle in core syllabification. In J. Kingston & M. E. Beckman (Eds.), *Papers in Laboratory Phonology I: Between the grammar and physics of speech* (pp. 283–333). Cambridge University Press.

Flowerdew, J. (1992). Research of Relevance to Second Language Lecture Comprehension: An Overview. *Perspectives*, *4*(2), 9–23.

Glyde, H., Cameron, S., Dillon, H., Hickson, L., & Seeto, M. (2013). The effects of hearing impairment and ageing on spatial processing. *Ear and Hearing*, *34*(1), 15–28.

Glyde, H., Hickson, L., Cameron, S., & Dillon, H. (2011). Problems hearing in noise in older adults. Spatial processing disorder? *Trends in Amplification*, *15*(3), 116–126.

Graydon, K., Rance, G., Dowell, R., & Van Dun, B. (2017). Consequences of early conductive hearing loss on long-term binaural processing. *Ear and Hearing, 38*(5), 621–627. doi:10.1097/AUD.0000000000000431

Hyman, L. (2007). Universals in phonolgy. *UC Berkeley Phonology Lab Annual Report*, 345–390. doi:10.2139/ssrn.1780284

Kilman, L., Zekveld, A., Hällgren, M., & Rönnberg, J. (2014). The influence of non-native language proficiency on speech perception performance. *Frontiers in Psychology*, *5*, 1–9. doi:10.3389/fpsyg.2014.00651

Krizman, J., Bradlow, A. R., Lam, S. S. Y., & Kraus, N. (2016). How bilinguals listen in noise: Linguistic and non-linguistic factors. *Bilingualism*, *20*(4), 834–843. doi:10.1017/S1366728916000444

Major, R. C., Fitzmaurice, S. F., Bunta, F., & Balasubramanian, C. (2002). The effects of nonnative accents on listening comprehension: Implications for ESL assessment. *TESOL Quarterly*, *36*(2), 173–190. doi:10.2307/3588329

Mealings, K., Cameron, S., Chong-White, N., Young, T., & Dillon, H. (in submission a) Listening in Spatialized Noise – Universal Test (LiSN-U) Test-retest reliability study. *International Journal of Audiology.*

Mealings, K., Cameron, S., & Dillon, H. (in submission b). Correlating performance on the Listening in Spatialized Noise – Sentences Test (LiSN-S) with the Listening in Spatialized Noise – Universal Test (LiSN-U). *International Journal of Audiology.*

Munro, M. J. (1998). The effects of noise of the intelligibility of foreign-accented speech. *SSLA*, *20*, 139–154. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=15816219

Phonemic Inventories and Cultural and Linguistic Information Across Languages. (n.d.). *American Speech-Language-Hearing Associated*. Retrieved January 25, 2019, from https://www.asha.org/practice/multicultural/Phono/

Rogers, C. L., Lister, J. J., Febo, D. M., Besing, J. M., & Abrams, H. B. (2006). Effects of bilingualism, noise, and reverberation on speech perception by listeners with normal hearing. *Applied Psycholinguistics*, *27*(3), 465–485. doi:10.1017/s014271640606036x

Schafer, E. C., Aoyama, K., Ho, T., Castillo, P., Conlin, J., Jones, J., & Thompson, S.

(2018). Speech recognition in noise in adults and children who speak English or Chinese as their first language. *Journal of the American Academy of Audiology*, *29*(10), 885–897. doi:10.3766/jaaa.17066

Skoe, E., & Karayanidi, K. (2018). Bilingualism and speech understanding in noise: Auditory and linguistic factors. *Journal of the American Academy of Audiology*, *30*(2), 115–130. doi:10.3766/jaaa.17082

Tabri, D., Chacra, K. M. S. A., & Pring, T. (2011). Speech perception in noise by monolingual, bilingual and trilingual listeners. *International Journal of Language and Communication Disorders*, *46*(4), 411–422. doi:10.3109/13682822.2010.519372

Tauroza, S., & Luk, J. (1997). Accent and second language listening comprehension. *RELC Journal*, *28*, 54–71.

Tomlin, D., & Rance, G. (2014). Long-term hearing deficits after childhood middle ear disease. *Ear and Hearing*, *35*(6), e233–e242. doi:10.1097/AUD.0000000000000065

Weinberger, S. H. (n.d.). *The speech accent archive*. Retrieved January 25, 2019, from http://accent.gmu.edu/browse_native.php

Tables

Table 4: Age and gender ANOVA results for spatially-separated, co-located, and spatial advantage measures.

		SS	*df*	*MS*	*F*	*p*
Spatially-separated	Intercept	7946	1	7946	1050	<0.0005
	Age	505	1	505	67	<0.0005
	Gender	9	1	9	1	0.29
	Error	1112	147	8		
Co-located	Intercept	463	1	463	267	<0.0005
	Age	35	1	35	20	<0.0005
	Gender	1	1	1	0.6	0.43
	Error	255	147	2		
Spatial advantage	Intercept	12229	1	12229	2103	<0.0005
	Age	277	1	277	48	<0.0005
	Gender	3	1	3	0.6	0.45
	Error	855	147	6		

Table 5: Analysis of residual for the spatially-separated, co-located, and spatial advantage data for 127 children. The correlation coefficient r shows the correlation of each measure with age.

Measure	*r*	*a*	*b*	*Var*	*SD*	*SD* of residual
Spatially-separated	-0.596	-6.2	-0.907	9.394	3.06	2.46
Co-located	-0.486	5.5	-0.347	1.934	1.39	1.22
Spatial advantage	0.433	11.7	0.560	6.955	2.64	2.38

Table 6: Means and standard deviations for spatially-separated, co-located, and spatial advantage data for 23 adults.

Measure	*Mean*	*SD*
Spatially-separated	-18.7	1.97
Co-located	1.3	0.89
Spatial advantage	20.0	2.07

Figure Captions

Figure 1: Graph showing mean SRTs on spatially-separated and co-located conditions, and spatial advantage measure by age group. Bars represent 95% confidence intervals.

Figure 3: Scatterplots of the a) spatially-separated, b) co-located, and c) spatial advantage scores in dB as a function of age for the 127 children in the present study.

Figure 3: LiSN-U and LiSN-S spatial advantage comparison for both adults and children. The dots represent the LiSN-U scores. The solid line represents the piece-wise linear regression for the LiSN-U (n = 150 from the present study). The dashed line represents the piece-wise linear regression from previously published normative data for the LiSN-S (n = 202).