

Evaluation of headphone effects on performance in the LiSN & Learn auditory training software

CHI YHUN LO,¹ HARVEY DILLON,² SHARON CAMERON,³
AND CATHERINE M. MCMAHON⁴

Abstract

Spatial processing is the ability to use binaural cues to separate and group sounds that are important for everyday listening. Spatial processing disorder (SPD) is the reduced ability to utilise binaural cues to segregate sounds. LiSN & Learn is auditory training software designed to remediate SPD, and in its current form can only be administered with a specific type of headphone (Sennheiser HD 215). The purpose of this study was to evaluate the effect that various headphones had on spatial processing ability in LiSN & Learn, with the aim of allowing flexibility of choice and increasing access to the program. Fourteen participants played training games across a counterbalanced test-retest design that measured speech reception thresholds (SRTs) across seven headphones. The results indicate the effect of the headphone was relatively minor, and performance in spatial processing had a strong correlation with the ability of the headphone to reproduce high frequencies. As such, it is recommended that any commercially available headphone with a high-frequency response broadly similar to a Sennheiser HD 215 is suitable for use with LiSN & Learn.

Introduction

Cherry (1957) first proposed the ‘Cocktail Party Problem’ to illustrate the capacity to listen to one speaker in the company of many competing voices. This ability to selectively attend to a target speaker and simultaneously suppress extraneous

1 Centre for Language Sciences, Department of Linguistics, Macquarie University.

2 National Acoustic Laboratories, Australian Hearing Hub, Macquarie University.

3 National Acoustic Laboratories, Australian Hearing Hub, Macquarie University.

4 Centre for Language Sciences, Department of Linguistics, Macquarie University.

noises requires the integration of spatial cues, and is known as spatial processing (Glyde et al., 2011). Two ears allow for binaural hearing, and any sound that arrives from outside the median plane (i.e. off-centre) will reach the ear closer to the sound source first, resulting in an interaural time difference (Palomäki et al., 2005). Interaural level differences will manifest due to the head itself acting as a masking barrier, attenuating high frequencies above about 1.5 kHz (Carlyon, 2004). An illustration of these two interaural differences is shown in Figure 1. Interaural time and level differences provide the basic mechanisms for sound localisation, but the primary benefit of binaural hearing in modern industrial society is more effective listening in noisy conditions (Moore, 1991). Binaural hearing allows sound to be segregated into streams – abstract cognitive percepts that are conducive to the organisation of sound, such as the separation between a target speaker and noise (Bregman, 1994).

Grouping and streaming sounds is fundamental for acoustic cohesion and contextual understanding, and is particularly important in the context of the classroom, where children spend 60 per cent to 75 per cent of their time engaged in listening tasks (Rosenberg, 2010). It is crucial that the students listen to their teacher and classmates when required, while ignoring background noise such as non-relevant student discussion or acoustic reverberation. Without this ability, the detrimental effect on educational performance may result in poorer psychological, emotional, and social outcomes (Musiek & Chermak, 2007).

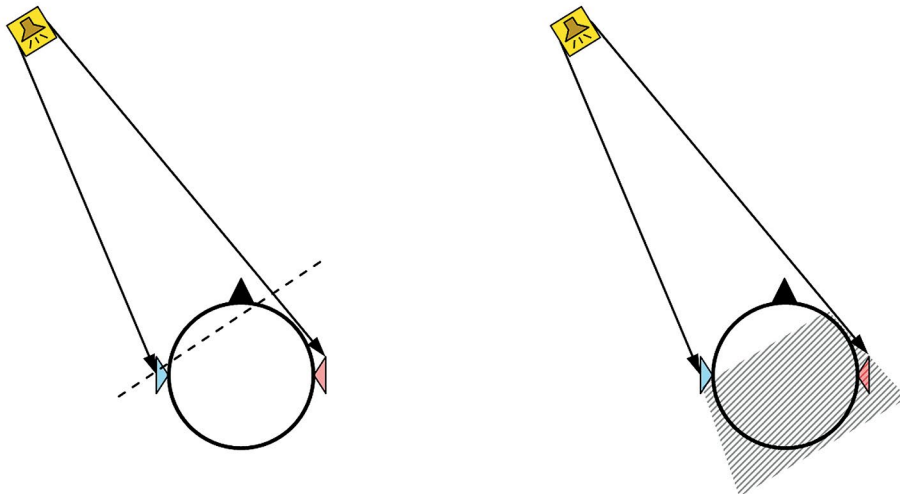


Figure 1. Illustration of interaural differences.

In the left diagram, the sound source will reach the left ear first as it is closer than the right ear, resulting in an interaural time difference. In the right diagram, the sound source will reach the left ear directly, but the head itself will act as a barrier and attenuate the sound source before it reaches the right ear, resulting in an interaural level difference.

Source: Chi Yhun Lo.

Spatial processing disorder (SPD) is the reduced ability to use interaural time and level differences to segregate sounds on the basis of direction, particularly in noisy environments, and falls under the umbrella term of 'central auditory processing disorder' (CAPD) (Cameron & Dillon, 2011). CAPD can be viewed as a heterogeneous disorder that refers to difficulty with auditory processes in the central auditory nervous system (CANS) not resulting from disorders of attention, cognition, language or peripheral hearing loss; though these may be (and commonly are) comorbid factors (ASHA, 2005; Dillon et al., 2012). The development of the CANS is critical until around 12 years of age, at which point it is mostly developed (Sanes & Woolley, 2011). Additionally, findings by Moore (2007) showed that children with previously impaired binaural abilities could learn those abilities at a later age than normal, underscoring our capacity for brain plasticity, and the importance of early diagnosis and remediation.

To overcome deficits from SPD, some forms of intervention include the use of assistive listening devices, improvement of room acoustics, or placement of the child towards the front of the class, closer to the teacher – all of which have the purpose of improving the signal-to-noise ratio (Miller et al., 2005). Cameron and Dillon (2008) developed the Listening in Spatialized Noise-Sentences Test (LiSN-S) as a diagnosis tool for SPD, and found that children with SPD required a far greater signal-to-noise ratio than their normal hearing counterparts (Cameron & Dillon, 2011).

Cameron and Dillon (2011) developed formal auditory training software designed to remediate SPD by improving binaural processing abilities in children. Designed for use in the home environment, LiSN & Learn requires the child to engage with training games through a pair of headphones on a personal computer. The speech recordings were convolved with head-related transfer functions (HRTFs, the sum of acoustic filtering effects that the pinna, head, and torso form), to create a virtual spatial environment. Even though the child is wearing headphones, the listener will perceive an externalised sound, equivalent to a free-field, three-dimensional environment (Cameron et al., 2012). For each game, the child must identify a target word nested within a sentence that is perceptually located directly in front (0° azimuth), while ignoring competing distractors from both left and right presentations ($\pm 90^\circ$ azimuth); only spatial cues can be used as both target and distractor sentences are presented by the same female speaker (Cameron & Dillon, 2011). An illustration of this spatial presentation can be seen in Figure 2. After the completion of 120 games, children with SPD showed an average improvement of 10.9 decibels (dB) for speech reception thresholds (SRTs, defined as the signal-to-noise ratio at which 50 per cent of words were correctly perceived), improving from -10.4 dB to -21.3 dB (Cameron et al., 2012); a lower score indicates better performance. Responses from parents, teachers, and self-reported questionnaires showed

positive outcomes with tangible, lasting benefits that are ecologically valid (Cameron et al., 2012). For a comprehensive review of the LiSN & Learn auditory training software see Cameron and Dillon (2013).

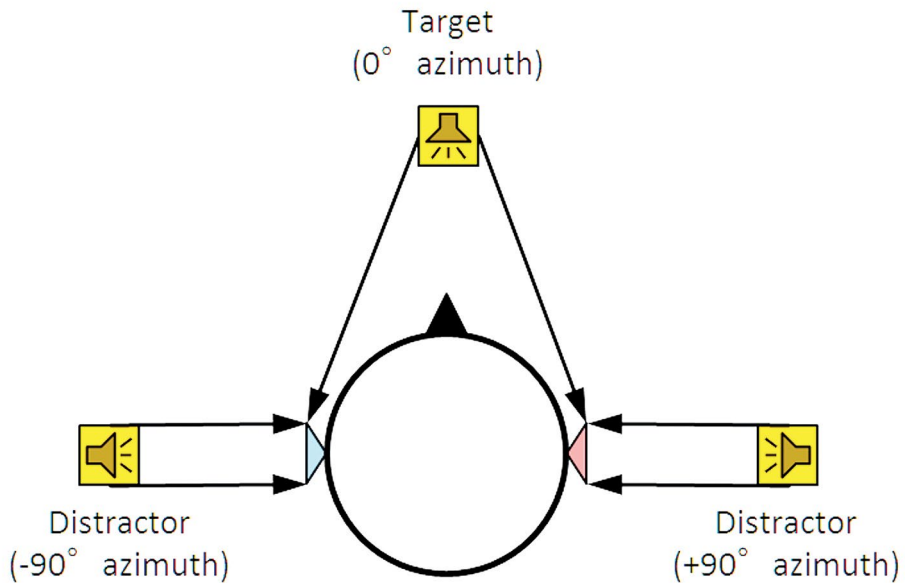


Figure 2. Illustration of spatial presentation.

For each game, the target speech is presented directly in front (0° azimuth) while the distractor speech is presented at left and right positions ($\pm 90^\circ$ azimuth).

Source: Chi Yhun Lo.

LiSN & Learn was specifically configured for use with Sennheiser HD 215 headphones that are currently included with the software package. The speech recordings were convolved with the inverse headphone-to-eardrum transfer function (HpTFs, the sum of acoustic filtering effects of the headphone) from a Sennheiser HD 215, as measured on a Knowles Electronics Manikin for Acoustic Research (KEMAR) manikin. The aim of this was to correct for the response of the headphone, removing additional spectral coloration, at least for people who have the same earphone-to-eardrum transfer function as that of KEMAR. However, this limits the test to being conducted with the Sennheiser HD 215 as the sole headphone for use with LiSN & Learn. This is problematic as headphones have a finite product life and production cycle, and even individual headphones within the same production run will have slightly different characteristics. Additionally, the limitation on choice is made more apparent when considering headphones are available in a myriad of types and styles, with differences in quality, design, application, and price ranges.

The requirement to use a particular headphone may be unnecessary. There is evidence that spatial processing with a front talker and two symmetrical distractors is primarily obtained through a process termed better-ear glimpsing, in which the brain extracts, at each moment in time and in each frequency range, signal components from whichever ear has the better signal-to-noise ratio (Brungart and Iyer, 2012; Glyde et al., 2013a), with these differences being caused by the target and distractors having different interaural level differences (Glyde et al., 2013b). If this were the only mechanism, the introduction of spectral coloration by headphones would have no effect, as at every frequency it would affect the target and distractors by exactly the same amount, and so leave the signal-to-noise ratio in each ear unaffected. But as the signal-to-noise ratio differences between the ears arise from head diffraction effects, and because head diffraction effects are much stronger for high frequencies than for low frequencies, achieving normal spatial processing is likely to rely on the headphones adequately reproducing the high-frequency components of speech.

This study examined seven headphone conditions to determine their efficacy in enabling speech intelligibility when the target and distractors come from different directions, as measured with the LiSN & Learn software, with the aim of allowing for flexibility of choice, potentially reducing economic barriers, and improving access to the program. It is hypothesised that headphone selection should not have an effect on LiSN & Learn SRTs provided the headphones adequately present the full audio bandwidth to the listener.

Method

Approval for this study was granted by the Macquarie University Faculty of Human Sciences Human Research Ethics Sub-Committee (reference: 5201300623), as well as the Australian Hearing Human Research Ethics Committee.

Headphone pilot study

Headphone classification is determined by their construction, and the IEC-Standard 60268-7 provides two broad distinctions (IEC, 1996). Earphone (EP): an electroacoustic transducer intended to be closely coupled acoustically to the ear; and headphone (HP): assembly of one or two earphones on a head/chin band. From these, our selection of headphones consisted of three sub-types. Intra-concha EP: a small earphone that fits in the concha cavity; supra-aural HP: a headphone applied externally to the outer ear and intended to rest on the pinna; circumaural HP: a headphone with a cavity large enough to enclose the entire ear.

Initially, 14 headphones were selected based on commercial and consumer availability, affordability, and across a wide range of common types (circumaural, supra-aural, and intra-concha). Headphones with active technology such as equalisation or noise-cancelling capability were excluded. Headphones were placed on (or in) the artificial ears of a Brüel & Kjær Head and Torso Simulator Type 4128C, with left and right output levels measured using two Brüel & Kjær Measuring Amplifiers Type 2610. Transfer functions were measured twice for reliability, using swept sine waves from 30 to 20,000 Hz with a Stanford Research Systems Model SR785 2-Channel Dynamic Signal Analyzer. Amplification to the headphones was provided by a Yamaha AX-350 set at a constant output of approximately 35 mV rms that provided peak output levels between approximately 85 and 95 dB across all headphones.

In addition to the Sennheiser HD 215 (circumaural HP) that is currently included with LiSN & Learn, five additional headphones were selected from the original pool of 14 headphones, with an emphasis on minimising cost to consumer, testing a wide range of characteristics, and with the following considerations: Apple EarPods (intra-concha EP) were selected based on their ubiquity and strong high-frequency response; the Koss KPH7 (supra-aural HP) is a low-cost headphone that had a similar transfer function to a more expensive and comparable supra-aural headphone; the Sennheiser HD 201 (circumaural HP) is a low-cost headphone with a similar transfer function to the higher cost Sennheiser HD 215; the Sony MDR-E9LP (intra-concha HP) has a low cost and relatively flat high-frequency response; and the Sony MDR-MA100 (circumaural HP) was selected as it had a sharp dip in response at 4 kHz, and strong fluctuating response (peaks and valleys) above 2,000 Hz. The headphones are characterised by their transfer functions from 30 to 20,000 Hz, and are presented in Figure 3.

Participants

Fourteen undergraduate students from Macquarie University volunteered for this study and were compensated for their time and participation. The students (eight female and six male) ranged in age from 19 to 24 ($M = 21$ years) with Australian English as their first language. On each day of testing, all participants showed normal hearing (≤ 20 dB Hearing Level) from 500 to 8,000 Hz.

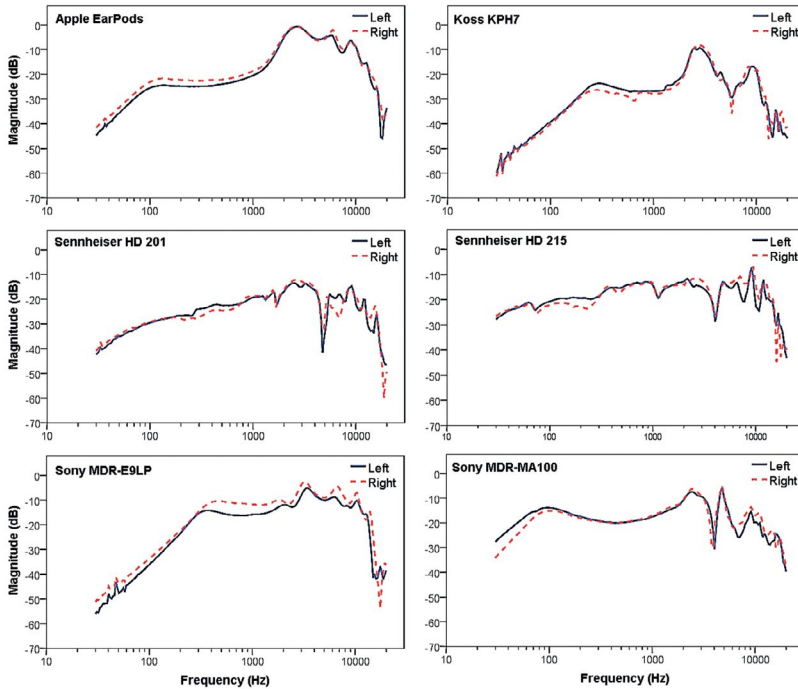


Figure 3. Headphone transfer functions for six headphones measured from 30 to 20,000 Hz.

Source: Chi Yhun Lo.

Materials

For each session, pure-tone audiometric screening was performed with a Maico MA 53 clinical audiometer using Sennheiser HAD 200 audiometric headphones. Two copies of LiSN & Learn software version 3.1.0 were used as test materials. One was a normal version (with the speech materials convolved with inverse Sennheiser HD 215 HpTF) installed on a Hewlett Packard Z200 Workstation desktop computer with only the Sennheiser HD 215 being used. The other was modified with the removal of the inverse HpTF (i.e. not designed specifically for use with the Sennheiser HD 215) and installed on a Lenovo T420s laptop computer, using the Apple EarPods, Koss KPH7, Sennheiser HD 201, Sennheiser HD 215, Sony MDRE9LP, and Sony MDR-MA100. LiSN & Learn consists of four games that differ only in graphical presentation but not in task: Answer Alley, Goal Game, Listening House, and Listening Ladder. An additional game called Space Maze was omitted from testing as it differs greatly in graphical user interface, test material, and instruction to the other four games. Participants were provided a copy of the icons and their corresponding word meaning prior to their session for the purpose of familiarisation. A copy was provided at each session and participants were permitted to refer to it at any time.

Procedures

Testing occurred in an acoustically treated test booth at the National Acoustic Laboratories. Participants played a practice game, followed by six LiSN & Learn training games, using a different set of headphones for each game. As each headphone had a different transfer function and output level, presentation levels were adjusted prior to each game. Participants were asked to adjust a slider until they could just perceive a whooshing sound (speech-shaped random noise); the distracting sounds were presented 40 dB above this level. Participants completed two sessions (session 1 and session 2) that occurred at the same time of day separated by a week, to achieve high test-retest reliability. The order of the headphone conditions was counterbalanced while each game was randomly allocated.

For each game, the participant was required to identify a target word nested within a sentence that was perceptually located directly in front (0° azimuth), while ignoring competing distractors from both left and right presentations ($\pm 90^\circ$ azimuth). An example target sentence would be 'The turtle played in the street', whereby the participant would be marked correct if they selected the corresponding turtle icon. The LiSN & Learn interface can be seen in Figure 4.



Figure 4. LiSN & Learn interface for Goal Game.

Source: Sharon Cameron and Harvey Dillon.

Results

Analysis was performed with SPSS Statistics version 22. Bar graphs of the seven headphones for both session 1 and session 2 are shown in Figure 5. Headphones have been organised according to SRTs averaged across both sessions, ranked from best to worst. The best-performing headphone was the Sennheiser HD 215 (with inverse HpTF) with SRTs of -23.3 dB for session 1 and -23.8 dB for session 2, while the worst-performing headphone was the Koss KPH7 with SRTs of -20.6 dB for session 1 and -20.8 dB for session 2.

A repeated-measures general linear model showed a small but non-significant improvement in SRT scores across sessions of 0.5 dB [$F(1,6) = 3.3, p = 0.092$], and a significant effect for the headphones [$F(1,6) = 5.7, p < 0.0001$]. Pairwise contrasts with a Bonferroni adjustment for multiple comparisons between each headphone indicated significance was isolated between the Koss KPH7 and three other headphones (Sennheiser HD 215 (with inverse HpTF), $p = 0.009$; Sennheiser HD 215, $p = 0.036$; and Sennheiser HD 201, $p = 0.017$).

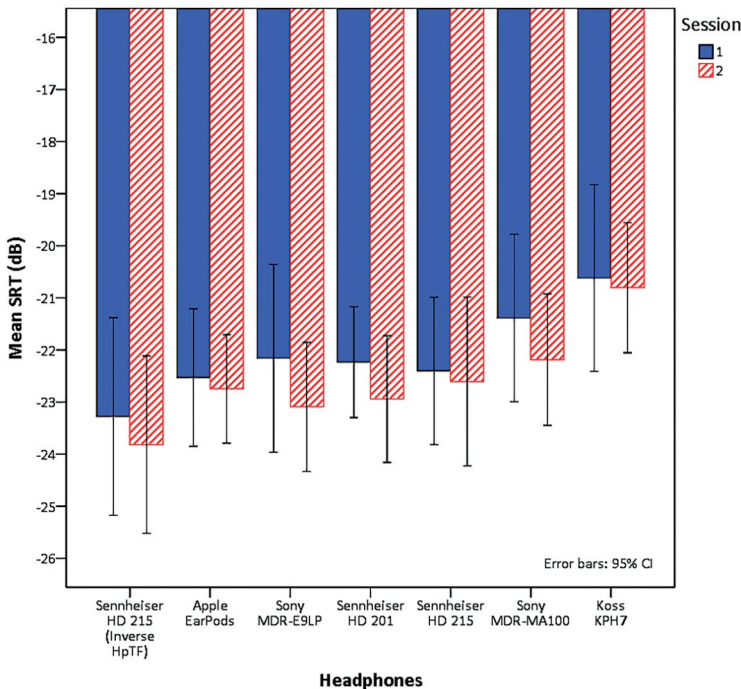


Figure 5. Bar graph of speech reception threshold (SRTs) for each headphone condition across session 1 and session 2.

Headphone conditions are ordered according to mean SRTs averaged across both sessions, from best to worst. Error bars represent 95 per cent confidence intervals. A lower SRT indicates better performance.

Source: Chi Yhun Lo.

An analysis of the headphone transfer functions was initiated on the premise that high frequencies may be a contributing factor for the Koss KPH7's low SRTs. The assumption was made that headphone transfer functions with a greater relative proportion of high frequencies would benefit spatial processing. Each headphone's relative high-frequency response was measured as the ratio of the average transfer function (in dB) between the frequencies from 8,000 to 20,000 Hz relative to its peak response. This particular frequency range was determined after visual examination of the differences between the measured responses of the headphones. Figure 6 shows a scatterplot of relative high-frequency response against mean SRTs. A one-tailed analysis showed a significant correlation with a strong positive linear association ($p = 0.044$, $r = 0.75$). The high correlation coefficient is, however, largely caused by the Koss KPH7 headphones, which had the weakest high-frequency response and the poorest SRT.

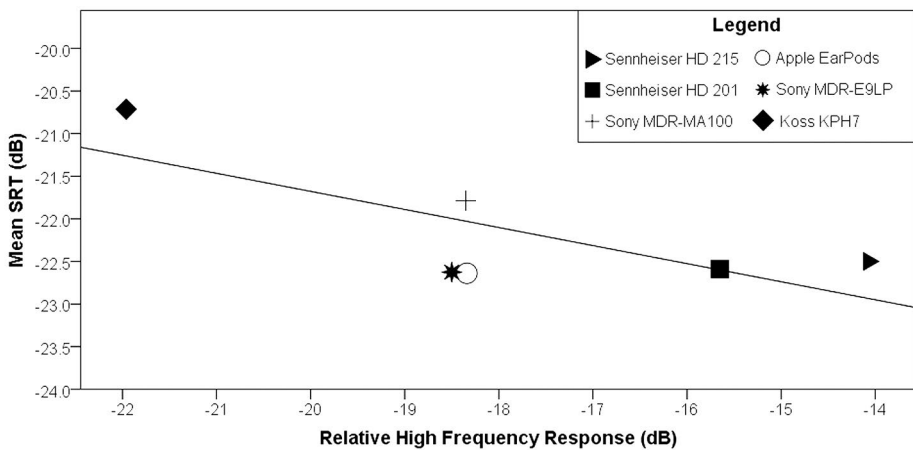


Figure 6. Correlation between headphone relative high-frequency response and spatial processing ability.

Mean speech reception thresholds (SRTs) across session 1 and 2 were plotted as a function of their relative high-frequency response, showing a significant correlation with a strong positive linear association ($p = 0.044$, $r = 0.75$).

Source: Chi Yhun Lo.

Discussion

The main aim of this study was to determine the effect of headphones on LiSN & Learn performance. If any headphone can be used, it allows for the flexibility of choice and improves access to LiSN & Learn. Overall, the choice of headphone did not have a significant effect on spatial processing performance, with the exception of the Koss KPH7. Each headphone's transfer function was

characterised as the ratio of its response from 8,000 to 20,000 Hz relative to its peak response. In an experiment on the influence of high frequencies on speech localisation, Best et al. (2005) found evidence that these frequencies were beneficial, and a potentially optimising feature for accurate spatial perception. This characterisation of the transfer functions showed three distinct groups: the Sennheiser HD 215 and Sennheiser HD 201 were the best performers with relative high-frequency responses at -14.1 dB and -15.7 dB respectively; the Apple EarPods, Sony MDR-MA100 and Sony MDR-E9LP were middle-range performers with relative high-frequency responses clustered around -18 dB; while the Koss KPH7 had the least relative high-frequency response at -22 dB. Figure 6 shows these results, providing an indication that the significant difference found in performance between the Koss KPH7 compared against the Sennheiser HD 201, Sennheiser HD 215, and Sennheiser HD 215 (with inverse HpTF) can be attributed to the relative reduction of high-frequency response in the Koss KPH7. The speech materials used in LiSN & Learn were digitally recorded in 16-bit and sampled at 44.1 kHz, so any lack of high frequencies can be attributed solely to the headphone response.

The strong positive linear association between relative high-frequency response and SRTs as shown in Figure 6 suggests that the high-frequency reproduction of a headphone has a significant effect on spatial processing performance, but this finding would need to be replicated across a wider range of headphones, and with a larger sample size, before it could be regarded as a reliable generalisation. The results in the present study support findings by Best et al. (2005), suggesting that high frequencies are an important feature of spatial processing. The results also provide further evidence that high frequencies from 8,000 to 20,000 Hz may be an important component of interaural level differences, and of the resulting interaural differences in SNR, and consequently the understanding of speech in spatially separated conditions.

Interestingly, the best-performing headphone condition was the Sennheiser HD 215 used in conjunction with speech materials convolved with the inverse HpTF, though the differences between this and most of the other headphones were small. The original decision to convolve the LiSN & Learn speech materials with the inverse HpTF was based on the assumption that it would minimise spectral coloration caused by the headphone, thus providing the listener with the most natural spectral cues. This appears to be the case, but the disadvantage of using the inverse HpTF is the limitation of the Sennheiser HD 215 as the sole headphone for use with LiSN & Learn. This study indicates that the effect on SRT when not correcting for the individual headphone transfer function is relatively minor (less than 2 dB when based on mean value and less than 1 dB when based on median values) for five out of the six earphones evaluated. Based on the results from this study, we conclude that any commercially available headphone

with relative high-frequency response broadly similar to the Sennheiser HD 215 is suitable for use with the LiSN & Learn auditory training software when the inverse HpTF is not present. Future iterations of the software will not convolve speech materials with an inverse HpTF.

Acknowledgements

We thank Kate Lloyd and Kath McLachlan at the PACE program of Macquarie University, for their support initiating this project. The authors would also like to thank Mark Seeto for his statistical advice.

Bibliography

- American Speech-Language-Hearing Association (ASHA). (2005). (Central) auditory processing disorders [Technical Report]. Available from www.asha.org/policy.
- Best, V., Carlile, S., Jin, C., and van Schaik, A. (2005). 'The role of high frequencies in speech localization,' *J. Acoust. Soc. Am.* 118, 353–363.
- Bregman, A. S. (1994). *Auditory Scene Analysis: The Perceptual Organization of Sound* (The MIT Press, Massachusetts).
- Brungart, D. S., and Iyer, N. (2012). 'Better-ear glimpsing efficiency with symmetrically placed interfering talkers,' *J. Acoust. Soc. Am.* 132, 2545–2556.
- Cameron, S., and 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,' *J. Am. Acad. Audiol.* 19, 377–391.
- Cameron, S., and 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,' *J. Am. Acad. Audiol.* 22, 678–96.
- Cameron, S., Glyde, H., and Dillon, H. (2012). 'Efficacy of the LiSN & Learn auditory training software: randomized blinded controlled study,' *Audiol. Res.* 2, 86–93.
- Cameron, S., and Dillon, H. (2013). 'Remediation of spatial processing issues in CAPD,' in *Handbook of Central Auditory Processing Disorders. Comprehensive Intervention* (Vol. II, pp. 201–224), edited by G. D. Chermak, and Frank E. Musiek (Plural Publishing, San Diego).

- Carlyon, R. P. (2004). 'How the brain separates sounds,' *Trends Cogn. Sci.* 8, 465–471.
- Cherry, C. (1957). *On human communication* (Wiley, New York).
- Dillon, H., Cameron, S., Glyde, H., Wilson, W., and Tomlin, D. (2012). 'An opinion on the assessment of people who may have an auditory processing disorder,' *J. Am. Acad. Audiol.* 23, 97–105.
- Glyde H., Buchholz J., Dillon H., Best V., Hickson L. and Cameron S. (2013a). 'The effect of better-ear glimpsing on spatial release from masking,' *J. Acoust. Soc. Am.* 134, 2937–2945.
- Glyde H., Buchholz J., Dillon H., Cameron S. and Hickson L. (2013b). 'The importance of interaural time differences and level differences in spatial release from masking,' *J. Acoust. Soc. Am.* 134, EL147-EL152.
- Glyde, H., Hickson, L., Cameron, S., & Dillon, H. (2011). 'Problems Hearing in Noise in Older Adults,' *Trends Amplif.* 15, 116–126.
- IEC. (1996). IEC 60268-7: Sound System equipment - Part 7: Headphones and Earphones .
- Miller, C.A., Uhring, E.A., Brown, J.J.C., Kowalski, E.M., Roberts, B., and Schaefer, B.A. (2005). 'Case studies of auditory training for children with auditory processing difficulties: a preliminary analysis,' *Contemp. Issues Commun. Science Disord.* 32, 93–107.
- Moore, D.R. (1991). 'Anatomy and Physiology of Binaural Hearing,' *Audiology.* 30, 125–134.
- Moore, D.R. (2007). 'Auditory processing disorders: acquisition and treatment,' *J. Commun. Disord.* 40, 295–304.
- Musiek, F.E., and Chermak, G.D. (Eds). (2007). *Handbook of (Central) Auditory Processing Disorder: Vol. 1. Auditory Neuroscience and Diagnosis* (Plural Publishing, San Diego).
- Palomäki, K.J., Tiitinen, H., Mäkinen, V., May, P.J.C., and Alku, P. (2005). 'Spatial processing in human auditory cortex: the effects of 3D, ITD, and ILD stimulation techniques,' *Cognitive Brain Res.* 24, 364–379.
- Rosenberg, G. (2010). 'Classroom Acoustics,' *Semin. Hearing,* 31, 188–202.
- Sanes, D.H., and Woolley, S.M.N. (2011). 'A behavioral framework to guide research on central auditory development and plasticity,' *Neuron,* 72, 912–929.

This text is taken from *The ANU Undergraduate Research Journal*,
Volume Six, 2014, edited by Jonathon Zapasnik and Alexandra Hogan,
published 2015 by ANU eView, The Australian National University,
Canberra, Australia.