A dynamic auditory-cognitive system supports speech-in-noise perception in older adults

Samira Anderson[a,b,*,e], **Travis White-Schwoch**[a,b], **Alexandra Parbery-Clark**[a,b], and **Nina Kraus**[a,b,c,d]

[a]Auditory Neuroscience Laboratory, Northwestern University, Evanston, IL 60208

[b]Department of Communication Sciences, Northwestern University, Evanston, IL 60208

[c]Department of Neurobiology and Physiology, Northwestern University, Evanston, IL 60208

[d]Department of Otolaryngology, Northwestern University, Evanston, IL 60208

Abstract

Understanding speech in noise is one of the most complex activities encountered in everyday life, relying on peripheral hearing, central auditory processing, and cognition. These abilities decline with age, and so older adults are often frustrated by a reduced ability to communicate effectively in noisy environments. Many studies have examined these factors independently; in the last decade, however, the idea of the auditory-cognitive system has emerged, recognizing the need to consider the processing of complex sounds in the context of dynamic neural circuits. Here, we use structural equation modeling to evaluate interacting contributions of peripheral hearing, central processing, cognitive ability, and life experiences to understanding speech in noise. We recruited 120 older adults (ages 55 to 79) and evaluated their peripheral hearing status, cognitive skills, and central processing. We also collected demographic measures of life experiences, such as physical activity, intellectual engagement, and musical training. In our model, central processing and cognitive function predicted a significant proportion of variance in the ability to understand speech in noise. To a lesser extent, life experience predicted hearing-in-noise ability through modulation of brainstem function. Peripheral hearing levels did not significantly contribute to the model. Previous musical experience modulated the relative contributions of cognitive ability and lifestyle factors to hearing in noise. Our models demonstrate the complex interactions required to hear in noise and the importance of targeting cognitive function, lifestyle, and central auditory processing in the management of individuals who are having difficulty hearing in noise.

Keywords

speech-in-noise perception; central processing; cognition; life experience; brainstem; auditory

*Corresponding author (CA): Samira Anderson, sander22@umd.edu, Hearing and Speech Sciences, University of Maryland, 0100 LeFrak Hall, College Park, MN 20742, USA, Tel: +1 301 405 4224, Fax: +1 301 314 2023.
[e]Now at Department of Hearing and Speech Sciences, University of Maryland, 0100 LeFrak Hall, College Park, MN 20742

I. Introduction

The ability to understand speech in noise involves a complex interplay of sensory and cognitive factors (Pichora-Fuller et al., 1995). This is especially true for older adults, who must compensate for the deterioration in peripheral sensory function that accompanies aging (Wong et al., 2010, Peelle et al., 2011). Three hypotheses have been proposed to explain the mechanisms underlying age-related deterioration in hearing in noise ability; they involve peripheral, central, and cognitive processes (CHABA, 1988, Humes, 1996) However, the age-related decline in speech-in-noise perception (SIN) is not inevitable; indeed, recent evidence demonstrates the beneficial effects of life-long experience (i.e., playing a musical instrument) in offsetting age-related effects on speech-in-noise perception in older adults (Parbery-Clark et al., 2011, Zendel and Alain, 2012). Even in the absence of intensive perceptual training, there is evidence for wide variability in speech-in-noise performance in older adults (McCoy et al., 2005, Harris et al., 2009, Schvartz et al., 2008). To understand this variability, we must consider the relative contributions of cognitive, central, and peripheral functions to hearing in noise in the aging auditory system.

1.1 Aging auditory system

Aging may affect speech-in-noise perception and central auditory processing in individuals with and without hearing loss, indicating the possibility that aging affects speech processing even when the signal is audible (Grose and Mamo, 2010, Gordon-Salant et al., 2006, Gordon-Salant and Fitzgibbons, 1993, Dubno, 1984, Poth et al., 2001, Harris et al., 2010, Vander Werff and Burns, 2011). In a meta-analysis Humes et al. (2012) concluded that age-related speech comprehension deficits arise from decreased peripheral and cognitive function. Peripheral and cognitive factors can affect performance on tests of central auditory processing, but there is no evidence in the perceptual literature to support declines that are specific to central processing itself. The studies examined by the meta-analysis all used behavioral measures of central auditory processing. While documentation of behavioral performance is important, it is also useful to consider aging effects on objective neural measures of speech. Indeed, studies documenting biological processing deficits have led to better understanding of the impairments in spectral and temporal processing in older adults (Anderson et al., 2012, Eckert et al., 2012, Ruggles et al., 2012, Schatteman et al., 2008). To date, however, they have not directly assessed the complex interactions among sensory, cognitive, and lifestyle factors that are in play when listening in noise.

Wingfield and colleagues (2005) argue that peripheral, central, and cognitive factors should not be examined independently but instead as a dynamic, integrated system to fully understand the difficulties encountered by older adults when listening in noise. This kind of study was undertaken by Humes et al. (1994) for older individuals (ages 63 to 83). Using canonical analysis, they found that hearing loss accounted for the largest share of variance in speech recognition measures, but cognition function also contributed to speech recognition in noise. Here, we aimed to expand on aspects of the Humes et al. study by including life experience factors and neural measures of speech processing, modeling speech-in-noise perception vis-à-vis recent theories of an integrated auditory-cognitive system (Holt and Lotto, 2008).

1.2 Auditory object: Role of cognition

Deciphering sounds in noise begins by forming a representation of an auditory object, allowing the listener to identify the sound (Bregman, 1990, Shinn-Cunningham and Best, 2008). For example, identifying a sound as the voice of a particular friend relies on accurate representations of speech cues, such as spectral and spatial cues (Du et al., 2011). Once an individual uses these cues to lock onto a particular voice, he is better able to follow that

voice in the conversation, a process known as stream segregation (Bregman and Campbell, 1971). Recently, physiologic evidence for stream segregation has been found in the auditory cortex; neural activity is selectively synchronized to the talker being attended to within a complex stream of sound (Ding and Simon, 2012, Mesgarani and Chang, 2012). When peripheral hearing loss and/or impaired auditory processing prevents the formation of a clear auditory object, however, the listener must rely on cognitive skills to facilitate top-down strengthening of an auditory signal (Nahum et al., 2008, Shinn-Cunningham and Best, 2008, Wingfield et al., 2005). Older listeners, too, rely on cognitive resources to facilitate comprehension (Pichora-Fuller et al., 2006, Pichora-Fuller, 2008), as declines in central auditory processing may hamper their ability to form an auditory object—even if audibility is corrected. For example, older adults draw from the prefrontal cortex—specifically, areas associated with memory and attention—to compensate for slower speed of processing or lack of clarity in sensory input (Wong et al., 2009). In fact, the volume (left pars triangularis) and thickness (superior frontal gyrus) of these prefrontal areas is related to speech-in-noise performance, but only in older adults (Wong et al., 2010). In younger adults, functional imaging reveals activity limited to the superior temporal gyrus (Wong et al., 2010), an area involved in auditory processing (Zatorre and Belin, 2001), language processing (Friederici et al., 2000), perception of facial emotions (Radua et al., 2010) and social cognition (Bigler et al., 2007). Moreover, older adults use linguistic knowledge (i.e., semantic and syntactic context of the sentence) to compensate for deficits in speech-in-noise perception arising from reduced processing speed; however, failing fluid memory abilities limit the ability to take advantage of this knowledge (Wingfield, 1996). This difficulty is compounded by older adults' frequent belief that they have correctly heard information they have in fact misperceived, especially when relying on context (Rogers et al., 2012); for this reason they may not ask for repetition or clarification. Because older adults who exhibit this "false hearing" are also likely to exhibit "false seeing" (Jacoby et al., 2012) or "false memory" (Hay and Jacoby, 1999), Jacoby and colleagues suggested that this convergence reflects a general deficit in frontal-lobe function. Impairments in executive function (i.e., inhibitory control) may also affect hearing in noise such that older adults are more easily distracted by novel auditory or visual stimuli, reducing their ability to selectively focus on the words spoken by a single speaker (Andrés et al., 2006).

1.3 Life experience

Although numerous studies have documented the effects of peripheral hearing loss, impaired central auditory processing, and reduced cognitive ability on speech-in-noise perception, relatively little is known about the modulating effects of life experiences. There is, however, some encouraging evidence that long-term life experiences affect the ability to hear in background noise. For example, musicianship offset the effects of aging on subcortical neural timing (Parbery-Clark et al., 2012) and is linked to better hearing in noise, auditory working memory, and temporal processing in older adults (Parbery-Clark et al., 2011, Zendel and Alain, 2012). High socioeconomic status, which is associated with higher levels of education and improved physical and cognitive health (Luo and Waite, 2005), may offset age-related declines in memory and other cognitive abilities (Czernochowski et al., 2008, Matilainen et al., 2010).

Life experiences which might modulate cognitive or neural systems are not limited to those that have occurred throughout the lifetime; current lifestyle activities can also play a role. For example, benefits of even a moderate amount of physical exercise are seen in overall cognitive health, including memory, attention, and executive function (Netz et al., 2011). The benefits of physical activity are also evident in brain structures; moderate aerobic exercise increases the size of the hippocampus (Erickson et al., 2011), the putative site of

episodic memory consolidation (Nadel and Moscovitch, 1997) and spatial navigation (Maguire et al., 2000).

Intellectual engagement may also benefit hearing in noise. Although the concept of "use it or lose it" as it relates to cognitive function is controversial (Swaab, 1991, Hultsch et al., 1999), positive associations between intellectual engagement and cognitive abilities have been demonstrated (as reviewed in Salthouse, 2006). For example, lifelong intellectual activity is associated with reduced levels of β-amyloid protein, a primary component of the amyloid plaques found in individuals with Alzheimer's disease (Landau et al., 2012). Engagement in activities involving high cognitive demands leads to a thicker cortical ribbon and increased neuronal density in Brodmann area 9 (Valenzuela et al., 2011), an area of the dorsolateral prefrontal cortex associated with inhibitory control in attentional selection (Dias et al., 1996), an important constituent mechanism of auditory stream segregation (Kerlin et al., 2010). Long-term engagement in these activities is best, but initiating these activities late in life benefits cognitive speed and performance consistency (Bielak et al., 2007, Smith et al., 2009). Taken together, these lifestyle factors may improve speech-in-noise perception skills by strengthening the compensatory mechanisms, especially the cognitive mechanisms that older adults rely on in the face of declines in peripheral and sensory processing.

1.4 Inferior colliculus—Top-down meets bottom-up

The auditory brainstem response to complex sounds (cABR), a far-field electrophysiological recording, offers a possible means for examining the reciprocal connections of sensory and cognitive circuits involved in speech-in-noise perception (as reviewed in Anderson and Kraus, 2010). The inferior colliculus (IC), the principle midbrain nucleus of the auditory pathway, is a site of intersection for numerous ascending and descending neural fibers (Ma and Suga, 2008) and is the putative generator of the cABR (Chandrasekaran and Kraus, 2010). The role of the IC in auditory learning has been demonstrated in animal models; for example, pharmacological disruption of the descending corticocollicular pathway prevents auditory learning in new circumstances (Bajo et al., 2010). In humans, the IC plays an important role in learning as well; in one study, Chandrasekaran et al. (2012) found that increased IC representation of pitch patterns, which distinguished words in a novel language, was associated with better perceptual learning. The cABR also reflects the modulating influences of both short- and long-term experiences on the nervous system, as evidenced by musical training (Bidelman et al., 2009, Wong et al., 2007, Skoe and Kraus, 2012) and language experience (Krishnan et al., 2008, Krishnan et al., 2005). Subcortical processing of speech in noise can also be modified by short-term (Song et al., 2012, Anderson et al., 2013b) and long-term training (Bidelman and Krishnan, 2010, Parbery-Clark et al., 2009).

Because IC activity reflects top-down modulation as well as local effects, such as age-related decreases in levels of inhibitory neurotransmitters (Caspary et al., 2005) and desynchronized neural firing (Hughes et al., 2010, Anderson et al., 2012), we expect that measures reflecting brainstem processing will contribute to a model of sensory-cognitive interactions in speech-in-noise perception. Given that the robustness of brainstem pitch and harmonic encoding (Anderson et al., 2011) and resistance to the degradative effects of noise (Parbery-Clark et al., 2009) are important factors in speech-in-noise perception, we have selected these measures of brainstem processing in our models to accompany peripheral and cognitive measures. We also expect that life experiences (exercise, intellectual engagement, socioeconomic status, and musicianship) influence the compensatory mechanisms that are engaged while listening in noise.

1.5 Current study

We hypothesized that speech-in-noise perception in middle- and older-aged adults relies on interactions of sensory, central, and cognitive factors, and that each of these factors makes distinct contributions both directly and indirectly through top-down modulation. We examined peripheral auditory function, cognitive ability, central auditory processing, and life experience in a large group of older adults. We took three approaches to testing our hypothesis, each involving a different statistical model of the dataset. In the first model, we used exploratory factor analysis to identify latent constructs (i.e., inferred underlying mechanisms) in our dataset. The factors were: (1) cognition (short-term memory, auditory working memory, auditory attention), (2) life experiences (intellectual engagement, physical activity, socioeconomic status), (3) central processing (brainstem measures of pitch, harmonics, and effects of noise), (4) peripheral hearing function (audiometric thresholds, otoacoustic emissions), and (5) speech-in-noise perception (three behavioral measures); these measures were therefore selected for subsequent models. In the second model we used structural equation modeling to do a confirmatory analysis of our hypotheses, determining the contributions of these factors, and their interactions to hearing in noise. In the third model we compared group differences in compensatory mechanisms for age-related decreases in speech-in-noise performance based on a history of musical training. We then used linear regression modeling to confirm the findings of structural equation modeling.

2. Methods

2.1. Participants

120 middle- and older-aged adults were recruited from the greater Chicago area (ages 55–79 years; mean 63.89, S.D. 4.83; 49 male). No participants had a history of a neurologic condition or hearing aid use. All participants had normal IQs (95 on the The Wechsler Abbreviated Scale of Intelligence; WASI; Zhu and Garcia, 1999), age-normal click-evoked auditory brainstem responses bilaterally (Wave V latency < 6.8 ms measured by a 100 µs click presented at 80 dB SPL at 31.25 Hz) and normal interaural Wave V latency differences (< 0.2 ms Hall, 2007).

2.2 Peripheral Function—Hearing

Audiometry—Air conduction thresholds were obtained bilaterally at octave intervals 0.125–8 kHz, with interoctave intervals at 3 and 6 kHz. All participants had pure tone averages (PTA; average threshold from .5–4 kHz) 45 dB HL and average thresholds from 6–8 kHz 80 dB HL. Thirty-nine participants (9 male) had normal hearing (defined as air conduction thresholds 20 dB HL at all frequencies). There were no asymmetries noted between the ears (> 15 dB HL difference at two or more frequencies) nor were there any air-bone conduction gaps (> 15 dB HL difference), meaning no conductive hearing losses were noted.

Otoacoustic emissions—Distortion product otoacoustic emissions (DPOAEs) were collected to measure cochlear outer hair cell function with SCOUT 3.45 (Bio-Logic Systems Corp., a Natus Company, Mundelein, IL). DPOAEs were assessed bilaterally from .75–8 kHz and we used the average amplitude of the distortion product across the frequency range as our measure.

2.3. Speech-in-noise perception measures

The QuickSIN is a non-adaptive clinical measure that is sensitive to performance differences between normal hearing and hearing impaired groups (Wilson et al., 2007). Four lists of six syntactically correct, low context sentences are presented binaurally at 70 dB HL through

insert earphones (ER-3A, Etymotic Research, Elk Grove Village, IL). The sentences, spoken by a female, are masked by a background noise of four-talker babble (three female, one male), with the first sentence in each set at a + 25 dB signal-to-noise ratio (SNR) and the SNR decreasing by 5 dB for each subsequent sentence down to a 0 dB SNR. Participants are asked to repeat each sentence as best they can. Each sentence contains five target words, and the participant scores a point for each one that they identify. The total number of words correctly repeated (maximum 30) is subtracted from 25.5 to obtain the SNR loss (dB), defined as the difference between an individual's speech-in-noise threshold and the average speech-in-noise threshold (Killion et al., 2004). The SNR loss scores are averaged across four lists to obtain the final SNR loss. Lower scores correspond to better hearing in noise.

The Words-in-Noise test (WIN) is a non-adaptive clinical measure (Wilson et al., 2007), chosen for its low memory and lexical demands. Participants are asked to repeat single words spoken by a female voice masked by four-talker babble (three female and one male) presented at 70 dB HL binaurally through insert earphones (ER-3A). Thirty-five words are presented starting at a 24 dB SNR and decreasing by 4 dB every five words until 0 dB SNR is reached. The final SNR score is based on the number of correctly-repeated words; lower scores indicate better performance.

The Hearing in Noise Test (HINT) is an adaptive clinical measure (Nilsson, 1994). Participants repeat short and semantically/syntactically simple sentences from the Bench-Kowal-Bamford sentence set (Bench et al., 1979) masked by speech-shaped background noise. Sentences are presented binaurally through insert earphones (ER-5A, Etymotic Research). The National Acoustics Laboratory-Revised (NAL-R; see *Stimulus*, below) amplification algorithm is applied in individuals with hearing loss. For all subjects, speech-shaped noise is set at 65 dB SPL and the intensity of the sentences is increased or decreased adaptively until a threshold SNR is determined, defined as the speech-noise level difference (in dB) at which the participant repeats 50% of the sentences correctly. A lower SNR indicates better performance.

2.4. Cognitive measures

The Auditory Attention Quotient (AAQ) from the Integrated Visual and Auditory Continuous Performance Test of Attention (IVA+, www.braintrain.com) was used to measure sustained attention. The IVA+ is a computer-based 20 minute "go - no go" test with 500 trials of either the number 1 (go cue) or 2 (no go cue) presented in a pseudorandom order in visual and auditory modalities; a practice run ensures that stimuli are viewable and audible. Responses are converted to age-normed standard scores. A higher score indicates better performance.

Auditory short-term memory was assessed by the Memory for Words subtest of the Woodcock-Johnson III Tests of Cognitive Ability (Woodcock et al., 2001). Participants repeat sequences of up to 7 words in the same order as the presentation. Words are presented by a CD at a comfortable intensity. Age-normed scores are used for analysis, and higher scores indicate better performance.

Auditory working memory was measured by the Auditory Working Memory subtest of the Woodcock-Johnson III Tests of Cognitive Ability. Participants are presented a series of 2 to 8 interleaved nouns (animals or food) and numbers, and are then instructed to repeat back the sequence by first parsing the nouns, and calling them out in their order of presentation, and then doing the same for the numbers. Stimuli are presented by a CD at a comfortable and audible intensity. Age-normed scores are used for analysis, and higher scores indicate better performance.

2.5. Life experience reports

Subjects self-reported life experiences pertinent to overall physical and cognitive health. Socioeconomic status (SES) was evaluated using educational criteria, a common metric which relates to other SES factors such as income (Farah et al., 2006, D'Angiulli et al., 2008). Subjects reported maternal and self-education levels on two Likert-type scales ranging from 1–4 (highest education level: middle school, high school/equivalent, college, or graduate/professional, respectively). The sum of these two scores was the SES level, and so scores ranged from 2–8. An intellectual engagement composite score was computed based on self-reported engagement in eight puzzle-like activities (crosswords, Sudoku, Scrabble™, "other word games," chess, reading, computer games, "other"). Each ranged from 0 (never) to 2 (regularly). The sum score was used, with scores ranging from 0 to 16. Finally, physical activity was evaluated with the General Practice Physical Activity Questionnaire (National Health Services, Department of Health, UK). Subjects rate their current levels of physical activity in employment and recreation. Scores are summed to produce a range from 0 to 19, with 19 indicating the highest level of physical activity.

2.6. Musical training

Subjects provided self-report histories of musical training, specifically: number of years of musical training, number of instruments played, age at start of musical training, etc. The three items were highly correlated and highly skewed – 52 of our participants answered 0 to these questions. Within those who reported a history of musical training, there was a severe positive skew as well, with only 7 reporting more than 40 years of musical training. For this reason, we created two groups based on years of musical training: one group (N = 52; Music −) had less than a year of musical training and the second group (N = 68; Music+) had a year or more of musical training (range: 1–71 years). Most of the musical training was early in life (during childhood). Seven people in the Music+ group played music professionally. The two groups were matched on all of the variables used in the model, except that the Music+ groups had higher scores for Socioeconomic Status, $t(119) = 2.686$, $p = 0.008$, Cohen's $d = 0.495$ and Short-Term Memory, $t(119) = 2.952$, $p = 0.002$, Cohen's $d = 0.544$ (Table 1).

2.7. Central Processing—Electrophysiology

Stimulus—A 170-ms speech syllable [da], with six formants, was synthesized in a Klatt-based formant synthesizer at a 20 kHz sampling rate. The stimulus is fully voiced except for an initial 5 ms stop burst, after which voicing remains constant with a 100 Hz fundamental frequency (F_0). The transition from the /d/ to the /a/ occurs during the first 50 ms of the syllable, and the lower three formants change linearly: F_1 400–720 Hz, F_2 1700–1240 Hz, and F_3 2580–2500 Hz. Formants are steady through the remaining 120 ms vowel portion. The remaining formants are constant throughout the stimulus: F_4 at 3300 Hz, F_5 at 3750 Hz, and F_6 at 4900 Hz. The waveform is presented in Figure 1A, with a spectrogram in Figure 1B.

The [da] was presented in quiet and noise conditions. In the quiet condition only the [da] is presented; in the noise condition masking is added at a +10 dB SNR. The masker was a six-talker babble background noise (3 female, English, 4000 ms, ramped; adapted from Van Engen and Bradlow, 2007), made by mixing 20 semantically-correct sentences. The babble track was looped continuously over the [da] to prevent phase synchrony between the speech stimulus and the noise.

In cases of hearing loss (thresholds > 20 dB HL at any frequency from .25–6 kHz for either ear), the stimulus was altered to equate audibility across subjects. This was achieved by selectively frequency-amplifying the [da] stimulus with the NAL-R algorithm (Byrne and

Dillon, 1986) using custom routines in MATLAB (The Mathworks Inc., Natick, MA) to create individualized stimuli. Our group has used this procedure previously and found that it improves the replicability and quality of the response (Anderson et al., 2013a). The presentation level of the background noise was adapted on a case-by-case basis to ensure that the hearing impaired participants, some of whom had stimuli presented at elevated SPLs, received a +10 dB SNR in the noise condition. During calibration, we determined the output of the amplified stimuli and then selected a babble noise file that had been amplified to achieve a +10 dB SNR.

Recording—Subcortical responses were recorded differentially, digitized at 20 kHz, using Neuroscan Acquire (Compumedics, Inc., Charlotte, NC) with electrodes in a vertical montage (Cz active, forehead ground, earlobe references; all impedances < 5 kΩ). The unamplified [da] syllable (i.e. stimulus for normal hearing individuals) was presented binaurally through electrically-shielded insert earphones (ER-3A; Etymotic Research) at 80 dB SPL with an 83 ms interstimulus interval in alternating polarities with Neuroscan Stim2. In cases of NAL-R-amplified stimuli (i.e., stimulus for individuals with hearing loss), the overall SPL was 80 dB or greater; however, all other parameters were identical.

The quiet condition always preceded the noise condition. Each condition lasted ~ 28 minutes, during which time participants sat in a recliner and watched a muted, captioned film to facilitate a relaxed, yet wakeful state. Six-thousand artifact-free sweeps of each response were collected (no more than 10% of the total sweeps were rejected due to artifact).

Data processing—Brainstem responses were digitally filtered offline, using a bandpass filter from 70–2000 Hz (Butterworth filter, 12 dB/octave roll-off, zero phase-shift) and epoched with a −40–213 ms time window referenced to stimulus onset. Artifact rejection was set at ± 35 μV. For analyses of noise degradation and pitch (see below, *Data Analysis*), responses to the two polarities were added, emphasizing the temporal envelope (Campbell et al., 2012, Gorga et al., 1985), whereas they were subtracted to emphasize stimulus fine structure for analyses of the first formant (Aiken and Picton, 2008). Final averages consisted of 6000 sweeps (3000 in each polarity) for each condition. Grand average waveforms in quiet and noise are presented in Figure 1C.

2.8. Data Analysis

Pitch & Formant—Fast Fourier transforms (FFTs) were run on brainstem responses to measure pitch (i.e., summed spectral amplitudes corresponding to the fundamental frequency and second harmonic) and formant (F_1) encoding (summed spectral amplitudes corresponding to the fourth through seventh harmonics) to the [da] in quiet. FFTs were run over the consonant-vowel transition (20–60 ms). Prior to the Fourier transform, zero-padding was applied to increase the resolution of the spectral display to 1 Hz/point; FFTs were then run with a Hanning window using a 4 ms ramp time. In each case, average spectral amplitudes were calculated over 60 Hz bins around frequencies of interest. Grand average FFTs are presented in Figure 2A and 2B.

Quiet-to-noise (Q-N) correlations—Responses in quiet and noise were cross-correlated to quantify objectively how much an individual is affected by the babble masking. The two responses are shifted in time (−2 to +2 ms) until the maximum correlation coefficient (Pearson product-moment correlation) is obtained. Maximum correlations are converted to Fischer *z* scores for statistical purposes (Cohen et al., 2003).

2.9. Statistical Analysis

Latent variable modeling—We used exploratory factor analysis and structural equation modeling to analyze the relative contributions of hearing, cognitive ability, central processing, and life experience to speech-in-noise perception. Model 1 presents the results of an exploratory factor analysis of our data, where we estimated latent factors modeled subsequently (Models 2 – 4). For these models, we used structural equation modeling, a latent variable modeling technique that considers the multiple covariances of observed items to identify, and elucidate relationships among latent variables (Anderson and Gerbing, 1988). The structural equation models follow a general confirmatory factor form which is described as follows in matrix algebra. Essentially, factor variances and covariances are estimated through linear regressions. The estimated covariances are then predicted by an overall factor model (Models 1 and 2).

For a given subject, s, the measurement model for each item, I, is estimated as follows:

$$X_{ls}=\mu_{ls}+\lambda_{ls}F_1+OF_2+OF_3+OF_4+OF_5+e_l$$
$$\dots OF_1+\lambda_{ls}F_s+OF_s+OF_4+OF_5$$
$$\dots$$
$$\dots OF_1+OF_2+OF_3+OF_4+\lambda_{ls}F_5 \quad (1)$$

where X is a given factor being estimated, λ is the unstandardized loading of regressing a given score Y on X, μ is the intercept or expected score Y when X is zero, F_{1-5} are the factor variances for each estimated latent variable, and e_I is the error variance for the estimated latent variable. This procedure is used for the one-factor model (Model 2).

The two-factor model (Model 3) follows the same general principles; and may be expressed overall to reflect group differences as follows:

$$S_j=\mu+\Lambda\eta_j+\varepsilon_j \quad (2)$$

Where S is performance on the speech in noise tasks, j refers to the categorical group (Music + or Music−), μ again to the regressor intercepts, Λ to the matrix of factor loadings, η to group differences in latent factors, and ε to standard errors (i.e., zero-mean residuals) for each group.

SEM-based procedures have several advantages over other statistical techniques such as principle components analysis or multiple linear regression. SEM provides the ability to model multiple predictor and criterion variables, construct latent variables (i.e, representations of unobserved qualities or conditions), include errors in modeling in a way that prevents bias from measurement error, and statistically test hypotheses regarding the interactions among multiple latent variables (Chin, 1998). Given outstanding questions about the reliance of speech-in-noise perception on peripheral, central, and cognitive function, then, it provides an ideal statistical modeling technique to explore the mechanisms underlying speech-in-noise perception.

Reliability estimates were computed on a factor level (Table 2). Three reliability estimates, ω_h ($omega_{hierarchical}$), ω_t ($omega_{total}$), and α are provided for each of the latent factors estimated and modeled (see sections 3.1 and 3.2). In all cases, higher coefficients refer to higher estimates of reliability. ω_h estimates the general factor saturation of a test scale, and essentially indicates what percentage of true score variance in a given factor the substituent items approximate (Zinbarg and Li, 2005). A variant, ω_t estimates the proportion of test variance that is due to all common factors (the "upper bound" estimate of reliability; Revelle and Zinbarg, 2009). Finally, we provide estimates of α, a familiar and popular coefficient of

reliability. While a problematic estimate of reliability or consistency of a scale (especially when attempting to identify one underlying factor), we report it nevertheless due to its ubiquity (see Streiner, 2003 for a detailed discussion).

The Shapiro-Wilk test for normality was used to ensure that all items were multivariate normal; log or root transforms were applied as appropriate to normalize items that were skewed. The exploratory factor analysis confirmed our selections of latent variables, in that they grouped together in the five factors proposed by our hypothesis. Factor loadings, or the variance of observed variables predicted by the 5 latent variables, are included in Table 3; see Model 1).

For each structural model, goodness of fit criteria included the χ^2 (chi-square) test and the root mean square error of approximation (RMSEA). The χ^2 test evaluates whether the model can be rejected ($p < 0.05$) or accepted by comparing the model to an optimal computer-generated model, thereby providing a formal test for over-identification of the model (i.e., an appropriately large number of samples relative to parameters); it is desirable to accept the null hypothesis of the χ^2 test. Specifically, the ratio of chi-square to degrees of freedom in the model should be approximately 1. The RMSEA is an *ad-hoc* measure of goodness of fit which estimates the fit of the model relative to a saturated model; values 0.05 are desirable. Data analyses were conducted in R (R Core Development Team), using the lavaan (Rosseel, 2012), GPArotation (Bernaards and Jennrich, 2005), and psych (Revelle, 2012) packages, and in SPSS 20.0 (SPSS, Inc., Chicago). Model selection for exploratory factor analysis was driven *a priori* by previous reports and evidence for correlations between specific factors and hearing in noise (Anderson et al., 2011); this analysis in turn informed subsequent confirmatory modeling. Only the single best-fit models are reported (Raykov et al., 1991). Refer to Figure 3 for symbol explanations in Figures 4 through 6.

Linear regressions were also performed, with speech-in-noise performance entered as the dependent variable and cognition, central processing, and life experience variables entered as the dependent variables. We used the "Enter" method of multiple linear regression to specify the order of variable entry. Linear regression modeling was then repeated with the Music+ and Music− groups. Collinearity diagnostics were run with satisfactory variance inflation factor (highest = 2.02) and tolerance (lowest = 0.494) scores, indicating the absence of strong correlations between two or more predictors. Table 4 presents inter-item correlations.

3. Results

3.1. Model 1: Identification of latent constructs

Exploratory factor analysis was applied to identify latent constructs used for subsequent modeling. A generalized least squares extraction method with varimax rotation and Kaiser normalization was used (Fabrigar and Wegener, 2011). Five factors were suggested by a scree test and estimated (see Table 3). Rotation converged with 6 iterations. The model was an appropriate fit, $\chi^2 = 24.433$, df = 40, $p = 0.975$; RMSEA = 0.0428. These factors (Hearing, Central Processing, Cognition, Life Experiences, and Hearing in Noise) were used in subsequent structural modeling (Models 2 and 3).

3.2. Models 2 and 3: Structural equation model of speech-in-noise perception

Model 2—Having identified latent variables, we examined their relative contributions to speech-in-noise performance. We used structural equation modeling which also allowed us to examine the interactions between constituent mechanisms of hearing in noise. The model converged normally after 63 iterations and was an appropriate fit, $\chi^2 = 72.59$, df = 67, $p =$

0.299; RMSEA = 0.035, accounting for 56.7% of the variance in speech-in-noise performance. In the model, the direct contributions of Central Processing and Cognition to Speech in Noise were significant; the contributions of Life Experiences and Hearing were not. Both Life Experiences and Cognition contributed to hearing in noise through top-down influences on central processing (Figure 4).

Model 3—We validated this model by performing a similar analysis on a random sample of 60% of the participants (N=72). Because of the reduced number of participants, we eliminated the variables that were not strong predictors to the original model or that incurred large error variance in the reduced subset (QuickSIN, Auditory Attention, Pitch Encoding, PTA, Physical Activity, and Intellectual Engagement). The model converged normally after 31 iterations and was an appropriate fit, $\chi^2 = 13.265$, df = 13, $p = 0.428$; RMSEA = 0.019, accounting for 36.4% of the variance in speech-in-noise performance. In this model, we again find that the Hearing variable does not significantly contribute to the model; however, Life Experiences is now a significant contributor, along with Cognition and Central Processing (Figure 5). Life Experiences and Cognition continue to contribute indirectly to speech-in-noise performance via modulation of Central Processing.

Post-hoc multiple linear regression was performed to confirm and replicate these findings, reducing the number of variables to six (Auditory Working and Short-term Memory, F1 Encoding, Q-N correlation, Physical Activity, and SES) – an appropriate number for a regression model with our sample size (Green, 1991). The HINT score was used as the dependent variable based on the strength of this variable in the SEM analysis. The independent variables were chosen based on the strength of their correlations with HINT and of their contributions to the SEM. Our model (F1 encoding, Q-N Correlation, Auditory Working Memory, Auditory Short-term Memory, SES, and Physical Activity) is a good fit for the data ($R^2 = 0.253$, $F(6,119) = 6.380$; $p < 0.001$). Only Auditory Working Memory and Q-N correlation, however, were significant contributors to the model. Table 5 provides standardized coefficients (β) and levels of significance for the predictors in the post-hoc regression.

3.3. Model 4: Two-group model based on history of musical experience

Because the music factor created two much smaller group models, the number of observed and latent variables needed to be reduced to assure that the model was not over-parameterized. Based on their minimal contributions as established in Model 2, Intellectual Engagement and Hearing were omitted *a posteriori* in Model 4. The model converged normally after 84 iterations and was an appropriate fit, $\chi^2 = 82.59$, df = 76, $p = 0.283$; RMSEA = 0.047. The χ^2 values for each group were 42.68 (Music+) and 39.91 (Music−). Our modeling indicates that the groups use different mechanisms for hearing in noise: in the Music+ group, Cognition was weighted more heavily (variance: Music+ = 0.54, Music− = 0.06), and in the Music− group the influence of Life Experience was stronger (variance: Music+ = 0.11, Music− = 0.71. Central Processing is an important factor for both groups (Figure 6).

To cross-validate the different weightings of Cognition and Life Experiences in the Music− and Music+ groups, we calculated 4 new sub-models and made the following comparisons: 1) a Music+ sub-model with Cognition but no Life Experiences factors vs. a Music+ sub-model with Life Experiences but no Cognition factors, and 2) a Music− sub-model with Life Experiences but no Cognition factors vs. a Music− sub-model with Cognition but no Life Experiences factors. All other variables were the same as those used in the original 2-group model. By comparing χ^2 differences between sub-models, we found that the Music+ sub-model with Cognition factors was a better fit than the one with Life Experiences factors (χ^2

= 15.276, df = 7, p = 0.033) and accounted for more variance in SIN perception (43% v 11%). In the Music− model, Life Experiences factors trended towards a better fit than the one with Cognition factors (χ^2 = 12.875, df = 7, p = 0.075) accounting for 79% vs. 35%. In addition, while Life Experiences did not significantly contribute to the Music+ sub-model with Life Experiences only, we found that Cognition factors were a significant contributor to the Music− sub-model with Cognition only. While our original model did not reveal the importance of Cognition factors for the Music− group, by cross-validating the model and removing Life Experiences, we see that Cognition factors do in fact play a role in speech-in-noise performance in both groups (Table 6).

We used multiple linear regression modeling to confirm these findings and to control for the effects of SES on the results. We entered "SES" on the first step and Auditory Working Memory, Q-N Correlation, and Physical Activity on the second step. On its own, SES did not significantly predict variance for either model (Music−: $F_{1,55}$ = 2.716, p = 0.105; Music +: $F_{1,63}$ = 0.116, p = 0.735). On the second step, both models are good fits for the data (Music−: $F_{1,55}$ = 3.973, p = 0.007; Music+: $F_{1,63}$ = 5.216, p = 0.001) with R^2 values of 0.238 (Music−) and 0.261 (Music+). Table 7 provides standardized coefficients (β) and levels of significance for the variables. In both of these models, only Auditory Working Memory significantly contributes to the variance on its own. The results of this model differed with those of our two-group SEM, suggesting that the SES factor may have been driving the results of the original two-group SEM.

4. Discussion

Our models reveal that cognition and central processing predict a significant proportion of the variance in speech-in-noise performance, with life experiences and cognition providing indirect influence by mediating the brainstem's effects through top-down modulation. In other words, cognitive abilities such as memory and attention, along with central (brainstem) processing of speech, help to determine how well an older individual understands speech in noise. Life experiences and cognition can also strengthen brainstem processing of sound, such that these factors also influence speech-in-noise perception indirectly. For the group that reported a history of musical training at any time in their lives (one year or more), cognition appears to play a more important role for understanding speech in noise, while life experiences play a bigger role in the group with no musical training. By controlling for SES, however, the role of cognition for the group with no musical training is also revealed. Central processing, as measured by the brainstem response to speech, predicts a significant amount of variance in both groups.

4.1. Auditory-cognitive system

The strong role of cognitive function in our model supports the view that the auditory system is dynamic and highly integrated with cognitive function (Arlinger et al., 2009). Cognition may be especially important for older adults as they draw on compensatory mechanisms to fill in gaps in an auditory stream caused by decreased audibility, slower sensory processing, and reduced spectrotemporal precision (Wong et al., 2009, Pichora-Fuller and Souza, 2003, Wong et al., 2010). The strong role of cognitive performance in our model is consistent with the Decline Compensation Hypothesis, which states that decreases in sensory processing can be compensated for, at least in part, by drawing on more general cognitive functions (Cabeza and Dennis, 2007). Interestingly, our models also suggest that this compensation can be modulated by previous life experience (musical training).

In our model, auditory working memory has the strongest cognitive influence, relative to auditory short-term memory and auditory attention, confirming a relationship between auditory working memory and speech-in-noise perception in older adults (McCoy et al.,

2005, Pichora-Fuller, 2003, Parbery-Clark et al., 2011). In a meta-analysis, Akeroyd (2008) found a fairly consistent relationship between speech recognition in noise and auditory working memory, while more general measures of IQ were not effective for predicting speech-in-noise ability. Others have found that speed of processing and attention are also predictive of speech-in-noise performance (Schvartz et al., 2008).

In addition to a direct effect, cognition may also have an indirect effect on speech-in-noise processing through modulation of brainstem processing. The Reverse Hierarchy Theory posits that cortical areas can access low-level information (i.e., brainstem processing) through top-down modulation to increase signal clarity, such as in the presence of background noise (Nahum et al., 2008). Shamma, Fritz and colleagues have demonstrated a neural basis for top-down modulation of sensory processing; they have established a functional connection between the frontal and auditory cortices in an animal model, such that cognitive function (attention) modulates sensory encoding in the auditory cortex (Fritz et al., 2010). A tentative connection between the frontal cortex and the brainstem has also been suggested by fMRI studies; for difficult tasks, the brainstem may signal the frontal cortex to increase attentional resources (Raizada and Poldrack, 2007).

4.2. Life experience

Life experience, comprising SES, physical activity, and intellectual engagement variables, contributed to speech-in-noise perception in our models, albeit less so than central processing or cognition. SES was the strongest of the items, and it likely accounted for variance that might otherwise have been observed in the physical activity or intellectual engagement variables. We used maternal and individual educational levels to determine SES; better education translates into better linguistic abilities that in turn aid the comprehension of spoken language (Wingfield and Tun, 2007). Benefits of higher SES for neural function have been established in children (Luo and Waite, 2005), and we conjecture that these benefits are sustained through the lifetime. Currently, adults approaching retirement age are seeking methods to mitigate the effects of aging. Given that cognitive engagement can reduce the manifestations of dementia through multiple pathways (Landau et al., 2012, Valenzuela et al., 2011), a healthy lifestyle comprising both physical and intellectual activities may reduce other expected effects of aging, such as impaired understanding of speech in noise.

4.3. Musical training and socioeconomic status

Although the two-group SEM suggested different weighting of Cognition versus Life Experience factors in individuals with and without musical experience, results of multiple linear regressions that control for SES indicate that Cognition is the most important variable in both groups. Therefore, SES appears to be a factor underlying the observed musician-nonmusician differences in preferential weighting of factors when listening to speech in noise. These results are not surprising, in that we would expect that families with greater resources (higher SES) would be more likely to enroll their children in musical training. Given that SES in the early years of life predicts cognitive function in later years, from adolescence to middle age (Osler et al., 2012), this early exposure to an enriched environment may facilitate reliance on cognitive resources seen in high SES groups.

Our results do not detract from the robust and compelling evidence that a life of musical training can change perceptual (Zendel and Alain, 2012, Parbery-Clark et al., 2011), cognitive (Chan et al., 1998, Cohen et al., 2011, Ho et al., 2003, Jakobson et al., 2008, Hanna-Pladdy and Gajewski, 2012, Tierney et al., 2008), and subcortical neural processing (Bidelman and Krishnan, 2010, Parbery-Clark et al., 2012, Wong et al., 2007); rather, they support the importance of considering SES in conjunction with musical training

(Schellenberg and Peretz, 2008, Schellenberg, 2005). Musical training requires active listening and engagement with sound and a connection of sound to meaning; for example, the dynamics of music (tempo, volume, etc.) can convey mood or imagery in such a way as to motivate a listener to carefully attend to an auditory stream (Kraus and Chandrasekaran, 2010, Patel, 2011). The use of music as a medium for enhanced educational and social outcomes has been successfully used in programs targeting children from lower SES backgrounds, such as El Sistema in Venezuela (Wakin, 2012) and the Harmony Project in Los Angeles (Gould, 2011), with these children achieving higher academic success than non-participating children (Majno, 2012). Therefore, we propose that although SES is a driving factor in the use of cognitive resources, musical training may be used to overcome the deficits that might result from an upbringing in a low SES background.

4.4. Hearing

Interestingly, hearing, as assessed by pure-tone audiometry and DPOAEs, did not make a significant contribution to the variance in speech-in-noise perception. These results are consistent with reports that the audiogram is not a good predictor of speech-in-noise performance (Killion and Niquette, 2000, Souza, 2007, Anderson et al., 2013a). It should be noted, however, that we excluded participants with more than a moderate hearing loss, so that average hearing levels were in the normal to mild hearing loss range. Had we included individuals with severe hearing loss, we expect that hearing would have accounted for more variance in speech-in-noise performance, as it did in the Humes et al (1994) study. The DPOAE measurement system we used (Scout 3.45), while providing a measure of the health of outer hair cells, cannot be used independently to predict hearing sensitivity or acuity (Kemp, 2002); therefore, it is not surprising that this measurement contributed minimally to the model. The suppression of DPOAEs through contralateral masking and the activation of the auditory efferent system may be related to speech-in-noise performance, but reports are conflicting. Some researchers have reported this relationship (de Boer et al., 2012, Mukari and Mamat, 2008), while at least one study found no relationship between contralateral suppression and speech-in-noise performance (Wagner et al., 2008). Future studies of sensory-cognitive relationships should consider a measure of efferent auditory system function, such as contralateral suppression of DPOAEs (Mazelová et al., 2003). Nevertheless, in our study two converging measures of peripheral auditory function failed to account for variance in speech-in-noise perception, further illustrating the critical roles of other factors.

4.5. Additional factors not included in the model

Although a sample size of 120 subjects is relatively large for electrophysiological studies, it is modest in the context of structural equation modeling. We are not making any claim that the measures we used are the sole factors involved in speech-in-noise perception; indeed, there are additional important factors which merit future investigation. It was important, however, that we exercise parsimony in selecting factors due to the constraints of our statistical modeling techniques.

For the cognitive factor, we included two measures of memory and one of attention, but not speed of processing. We acknowledge that speed of processing is an important factor in speech perception, especially in older adults who are affected by neural slowing and decreased temporal precision (Craig, 1992, Wingfield et al., 1999, Tremblay et al., 2002, Parthasarathy and Bartlett, 2011, Walton et al., 1998, Finlayson, 2002, Recanzone et al., 2011, Anderson et al., 2012) and future work will investigate what contributions it makes in an interactive speech-in-noise system.

We used the cABR as an objective measure of central auditory processing, but we could also have included additional electrophysiological measures, such as cortical evoked auditory potentials. Components of the cortical response waveform are related to speech-in-noise perception in children (Anderson et al., 2010), young adults (Parbery-Clark et al., 2011), and older adults (Getzmann and Falkenstein, 2011). Specifically, there is evidence that earlier components of the cortical evoked response (P1 and N2) reflect encoding acoustic cues such as frequency and timing whereas the later components (P2 and N2) reflect the synthesis of these features into a sensory representation (i.e., an auditory object) (Shtyrov et al., 1998, Ceponiene et al., 2005, Toscano et al., 2010). However, early cortical processing is closely related to brainstem processing (Abrams et al., 2006); therefore, we would predict that the addition of cortical measures would not contribute substantially more than brainstem processing. Nevertheless, future models of speech-in-noise perception should consider measures of cortical auditory processing, especially cortical components reflecting cognitive functions such as attention and working memory (Alho et al., 2012).

4.6. Clinical implications

The finding that central processing and cognition, but not hearing, were strong factors in our model speak to the need to address these issues when developing treatment plans for patients complaining of hearing-in-noise difficulties, including individuals with mild-to-moderate hearing loss. Developers of hearing aid algorithms have begun to consider cognitive function in hearing aid processing and have recommended slow-acting compression for individuals with low cognitive abilities (Lunner, 2003, Rudner et al., 2011, Cox and Xu, 2010, Gatehouse et al., 2003); in the future, engineers may consider how impaired central processing responds to different amplification algorithms. These factors can also be considered when designing training programs. There is evidence that central processing can be improved with short-term (Song et al., 2012, Russo et al., 2010, Carcagno and Plack, 2011, Hornickel et al., 2012, Anderson et al., 2013b) and long-term training (e.g. music or language, Parbery-Clark et al., 2009, Krizman et al., 2012, Bidelman and Krishnan, 2010, Krishnan et al., 2005, Parbery-Clark et al., 2012), and that training can improve cognition in older adults (Smith et al., 2009, Berry et al., 2010, Gazzaley et al., 2005). Our results suggest that perceptual training which incorporates high cognitive demands may improve speech-in-noise perception directly by training important factors such as memory and attention, and indirectly by strengthening cortical-subcortical sound-to-meaning relationships.

Acknowledgments

Abbreviations

cABR	auditory brainstem response to complex sounds
IC	inferior colliculus
IQ	intelligence quotient
DPOAE	distortion product otoacoustic emission
SNR	signal-to-noise ratio
SES	socioeconomic status

NAL-R	National Acoustics Laboratory-Revised amplification algorithm
SPL	sound pressure level
FFT	Fast Fourier Transform
Q-N	quiet-to-noise
SEM	structural equation modeling
RMSEA	root mean square error of approximation
PTA	pure tone average
SIN	speech in noise
fMRI	functional magnetic resonance imagining
SD	standard deviation
SE	standard error
Cog	Cognition

References

Abrams DA, Nicol T, Zecker SG, Kraus N. Auditory brainstem timing predicts cerebral asymmetry for speech. J Neurosci. 2006; 26:11131–11137. [PubMed: 17065453]

Aiken SJ, Picton TW. Envelope and spectral frequency-following responses to vowel sounds. Hear Res. 2008; 245:35–47. [PubMed: 18765275]

Akeroyd MA. Are individual differences in speech reception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults. Int J Audiol. 2008; 47:53–71.

Alho K, Grimm S, Mateo-León S, Costa-Faidella J, Escera C. Early processing of pitch in the human auditory system. Eur J Neurosci. 2012; 36:2972–2978. [PubMed: 22765015]

Anderson JC, Gerbing DW. Structural equation modeling in practice: A review and recommended two-step approach. Psychol Bull. 1988; 103:411–423.

Anderson S, Chandrasekaran B, Yi HG, Kraus N. Cortical-evoked potentials reflect speech-in-noise perception in children. Eur J Neurosci. 2010; 32:1407–1413. [PubMed: 20950282]

Anderson S, Kraus N. Sensory-cognitive interaction in the neural encoding of speech in noise: A review. J Am Acad Audiol. 2010; 21:575–585. [PubMed: 21241645]

Anderson S, Parbery-Clark A, White-Schwoch T, Kraus N. Aging affects neural precision of speech encoding. J Neurosci. 2012; 32:14156–14164. [PubMed: 23055485]

Anderson S, Parbery-Clark A, White-Schwoch T, Kraus N. Auditory brainstem response to complex sounds predicts self-reported speech-in-noise performance. J Speech Lang Hear Res. 2013a 1092–4388_2012_12-0043.

Anderson S, Parbery-Clark A, Yi HG, Kraus N. A neural basis of speech-in-noise perception in older adults. Ear Hear. 2011; 32:750–757. [PubMed: 21730859]

Anderson S, White-Schwoch T, Parbery-Clark A, Kraus N. Reversal of age-related neural timing delays with training. Proc Natl Acad Sci USA. 2013b

Andrés P, Parmentier FBR, Escera C. The effect of age on involuntary capture of attention by irrelevant sounds: A test of the frontal hypothesis of aging. Neuropsychologia. 2006; 44:2564–2568. [PubMed: 16797613]

Arlinger S, Lunner T, Lyxell B, Kathleen Pichora-Fuller M. The emergence of cognitive hearing science. Scand J Psychol. 2009; 50:371–384. [PubMed: 19778385]

Bajo VM, Nodal FR, Moore DR, King AJ. The descending corticocollicular pathway mediates learning-induced auditory plasticity. Nat Neurosci. 2010; 13:253–260. [PubMed: 20037578]

Bench J, Kowal A, Bamford J. The bkb (bamford-kowal-bench) sentence lists for partially-hearing children. Br J Audiol. 1979; 13:108–112. [PubMed: 486816]

Bernaards CA, Jennrich RI. Gradient projection algorithms and software for arbitrary rotation criteria in factor analysis. Educ Psychol Meas. 2005; 65:676–696.

Berry AS, Zanto TP, Clapp WC, Hardy JL, Delahunt PB, Mahncke HW, Gazzaley A. The influence of perceptual training on working memory in older adults. Plos ONE. 2010; 5:e11537. [PubMed: 20644719]

Bidelman GM, Gandour JT, Krishnan A. Cross-domain effects of music and language experience on the representation of pitch in the human auditory brainstem. J Cogn Neurosci. 2009; 23:425–434. [PubMed: 19925180]

Bidelman GM, Krishnan A. Effects of reverberation on brainstem representation of speech in musicians and non-musicians. Brain Res. 2010; 1355:112–125. [PubMed: 20691672]

Bielak, AaM; Hughes, TF.; Small, BJ.; Dixon, RA. It's never too late to engage in lifestyle activities: Significant concurrent but not change relationships between lifestyle activities and cognitive speed. J Gerontol B Psychol Sci Soc Sci. 2007; 62:331–339.

Bigler ED, Mortensen S, Neeley ES, Ozonoff S, Krasny L, Johnson M, Lu J, Provencal SL, Mcmahon W, Lainhart JE. Superior temporal gyrus, language function, and autism. Dev Neuropsychol. 2007; 31:217–238. [PubMed: 17488217]

Bregman, AS. Auditory scene analysis. Cambridge: MIT; 1990.

Bregman AS, Campbell J. Primary auditory stream segregation and perception of order in rapid sequences of tones. J Exp Psychol. 1971; 89:244–249. [PubMed: 5567132]

Byrne D, Dillon H. The national acoustic laboratories' (nal) new procedure for selecting the gain and frequency response of a hearing aid. Ear Hear. 1986; 7:257–265. [PubMed: 3743918]

Cabeza, R.; Dennis, N. Neuroimaging of healthy cognitive aging. In: CRAIK, FMI.; SALTHOUSE, TA., editors. The handbook of aging and cognition. 3. London: Psychology Press; 2007.

Campbell T, Kerlin J, Bishop C, Miller L. Methods to eliminate stimulus transduction artifact from insert earphones during electroencephalography. Ear Hear. 2012; 33:144–150. [PubMed: 21760513]

Carcagno S, Place C. Subcortical plasticity following perceptual learning in a pitch discrimination task. J Assoc Res Otolaryngol. 2011; 12:89–100. [PubMed: 20878201]

Caspary DM, Schatteman TA, Hughes LF. Age-related dhanges in the inhibitory response properties of dorsal cochlear nucleus output neurons: Role of inhibitory inputs. J Neurosci. 2005; 25:10952–10959. [PubMed: 16306408]

Ceponiene R, Alku P, Westerfield M, Torki M, Townsend J. Erps differentiate syllable and nonphonetic sound processing in children and adults. Psychophysiology. 2005; 42:391–406. [PubMed: 16008768]

Chan AS, Ho YC, Cheung MC. Music training improves verbal memory. Nature. 1998; 396:128–128. [PubMed: 9823892]

Chandrasekaran B, Kraus N. The scalp-recorded brainstem response to speech: Neural origins and plasticity. Psychophysiology. 2010; 47:236–246. [PubMed: 19824950]

Chandrasekaran B, Kraus N, Wong PCM. Human inferior colliculus activity relates to individual differences in spoken language learning. J Neurophysiol. 2012; 107:1325–1336. [PubMed: 22131377]

Chin WW. Commentary: Issues and opinion on structural equation modeling. MIS quarterly. 1998; 22:vii–xvi.

Cohen, J.; Cohen, P.; West, S.; Aiken, L. Applied multiple regression/correlation analysis for the behavioral sciences. 3. Mahwah, NJ: Erlbaum; 2003.

Cohen MA, Evans KK, Horowitz TS, Wolfe JM. Auditory and visual memory in musicians and nonmusicians. Psychon Bull Rev. 2011; 18:586–591. [PubMed: 21374094]

Cox RM, Xu J. Short and long compression release times: Speech understanding, real-world preferences, and association with cognitive ability. J Am Acad Audiol. 2010; 21:121–138. [PubMed: 20166313]

Craig CH. Effects of aging on time-gated isolated word-recognition performance. J Speech Hear Res. 1992; 35:234. [PubMed: 1735973]

Czernochowski D, Fabiani M, Friedman D. Use it or lose it? Ses mitigates age-related decline in a recency/recognition task. Neurobiol Aging. 2008; 29:945–958. [PubMed: 17280741]

D'angiulli A, Herdman A, Stapells D, Hertzman C. Children's event-related potentials of auditory selective attention vary with their socioeconomic status. Neuropsychology. 2008; 22:293–300. [PubMed: 18444707]

De Boer J, Thornton ARD, Krumbholz K. What is the role of the medial olivocochlear system in speech-in-noise processing? J Neurophysiol. 2012; 107:1301–1312. [PubMed: 22157117]

Dias R, Robbins T, Roberts A. Dissociation in prefrontal cortex of affective and attentional shifts. Nature. 1996; 380:69–72. [PubMed: 8598908]

Ding N, Simon JZ. Emergence of neural encoding of auditory objects while listening to competing speakers. Proc Natl Acad Sci, USA. 2012

Du Y, He Y, Ross B, Bardouille T, Wu X, Li L, Alain C. Human auditory cortex activity shows additive effects of spectral and spatial cues during speech segregation. Cereb Cortex. 2011; 21:698–707. [PubMed: 20685854]

Dubno J, Dirks D, Morgan D. Effects of age and mild hearing loss on speech recognition in noise. J Acoust Soc Am. 1984; 76:87–96. [PubMed: 6747116]

Erickson KI, Voss MW, Prakash RS, Basak C, Szabo A, Chaddock L, Kim JS, Heo S, Alves H, White SM, Wojcicki TR, Mailey E, Vieira VJ, Martin SA, Pence BD, Woods JA, Mcauley E, Kramer AF. Exercise training increases size of hippocampus and improves memory. Proc Natl Acad Sci USA. 2011; 108:3017–3022. [PubMed: 21282661]

Fabrigar, L.; Wegener, D. Exploratory factor analysis (understanding statistics). Cary, NC: Oxford University Press, USA; 2011.

Farah MJ, Shera DM, Savage JH, Betancourt L, Giannetta JM, Brodsky NL, Malmud EK, Hurt H. Childhood poverty: Specific associations with neurocognitive development. Brain Res. 2006; 1110:166–174. [PubMed: 16879809]

Finlayson PG. Paired-tone stimuli reveal reductions and alterations in temporal processing in inferior colliculus neurons of aged animals. J Assoc Res Otolaryngol. 2002; 3:321–331. [PubMed: 12382106]

Friederici AD, Meyer M, Von Cramon DY. Auditory language comprehension: An event-related fmri study on the processing of syntactic and lexical information. Brain Lang. 2000; 74:289–300. [PubMed: 10950920]

Fritz JB, David SV, Radtke-Schuller S, Yin P, Shamma SA. Adaptive, behaviorally gated, persistent encoding of task-relevant auditory information in ferret frontal cortex. Nat Neurosci. 2010; 13:1011–1019. [PubMed: 20622871]

Gatehouse S, Naylor G, Elberling C. Benefits from hearing aids in relation to the interaction between the user and the environment. Int J Audiol. 2003; 42:S77–S85. [PubMed: 12918613]

Gazzaley A, Cooney JW, Mcevoy K, Knight RT, D'esposito M. Top-down enhancement and suppression of the magnitude and speed of neural activity. J Cogn Neurosci. 2005; 17:507–517. [PubMed: 15814009]

Getzmann S, Falkenstein M. Understanding of spoken language under challenging listening conditions in younger and older listeners: A combined behavioral and electrophysiological study. Brain Res. 2011; 1415:8–22. [PubMed: 21880303]

Gordon-Salant S, Fitzgibbons PJ. Temporal factors and speech recognition performance in young and elderly listeners. J Speech Hear Res. 1993; 36:1276–1285. [PubMed: 8114494]

Gordon-Salant S, Yeni-Komshian GH, Fitzgibbons PJ, Barrett J. Age-related differences in identification and discrimination of temporal cues in speech segments. J Acoust Soc Am. 2006; 119:2455–2466. [PubMed: 16642858]

Gorga, M.; Abbas, P.; Worthington, D. Stimulus calibration in abr measurements. In: JACOBSEN, J., editor. The auditory brainstem response. San Diego: College Hill Press; 1985.

Gould, JE. Mozart vs. The gangstas: How classical music is changing young lives. Time U.S; 2011.

Green SB. How many subjects does it take to do a regression analysis? Multivariate Behavioral Research. 1991; 26:499–510.

Grose JH, Mamo SK. Processing of temporal fine structure as a function of age. Ear Hear. 2010; 31:755–760. [PubMed: 20592614]

Hall, J. New handbook of auditory evoked responses. Boston, MA: Allyn & Bacon; 2007.

Hanna-Pladdy B, Gajewski B. Recent and past musical activity predicts cognitive aging variability: Direct comparison with general lifestyle activities. Front Hum Neurosci. 2012; 6:1–11. [PubMed: 22279433]

Harris KC, Dubno JR, Keren NI, Ahlstrom JB, Eckert MA. Speech recognition in younger and older adults: A dependency on low-level auditory cortex. J Neurosci. 2009; 29:6078–6087. [PubMed: 19439585]

Harris KC, Eckert MA, Ahlstrom JB, Dubno JR. Age-related differences in gap detection: Effects of task difficulty and cognitive ability. Hear Res. 2010; 264:21–29. [PubMed: 19800958]

Hay JF, Jacoby LL. Separating habit and recollection in young and older adults: Effects of elaborative processing and distinctiveness. Psychol Aging. 1999; 14:122. [PubMed: 10224637]

Ho YC, Cheung MC, Chan AS. Music training improves verbal but not visual memory: Cross-sectional and longitudinal explorations in children. Neuropsychology. 2003; 17:439–450. [PubMed: 12959510]

Holt LL, Lotto AJ. Speech perception within an auditory cognitive science framework. Curr Dir Psychol Sci. 2008; 17:42–46. [PubMed: 19060961]

Hornickel J, Zecker SG, Bradlow AR, Kraus N. Assistive listening devices drive neuroplasticity in children with dyslexia. Proc Natl Acad Sci USA. 2012; 109:16731–16736. [PubMed: 22949632]

Hughes LF, Turner JG, Parrish JL, Caspary DM. Processing of broadband stimuli across a1 layers in young and aged rats. Hear Res. 2010; 264:79–85. [PubMed: 19772906]

Hultsch DF, Hertzog C, Small BJ, Dixon RA. Use it or lose it: Engaged lifestyle as a buffer of cognitive decline in aging? Psychol Aging. 1999; 14:245. [PubMed: 10403712]

Humes L. Speech understanding in the elderly. J Am Acad Audiol. 1996; 7:161–167. [PubMed: 8780988]

Humes LE, Watson BU, Christensen LA, Cokely CG, Halling DC, Lee L. Factors associated with individual differences in clinical measures of speech recognition among the elderly. J Speech Hear Res. 1994; 37:465–474. [PubMed: 8028328]

Jacoby LL, Rogers CS, Bishara AJ, Shimizu Y. Mistaking the recent past for the present: False seeing by older adults. Psychol Aging. 2012; 27:22. [PubMed: 22023508]

Jakobson LS, Lewycky ST, Kilgour AR, Stoesz BM. Memory for verbal and visual material in highly trained musicians. Music Percept. 2008; 26:41–55.

Kemp DT. Otoacoustic emissions, their origin in cochlear function, and use. Brit M Bull. 2002; 63:223–241. [PubMed: 12324396]

Kerlin JR, Shahin AJ, Miller LM. Attentional gain control of ongoing cortical speech representations in a "cocktail party". J Neurosci. 2010; 30:620–628. [PubMed: 20071526]

Killion M, Niquette P. What can the pure-tone audiogram tell us about a patient's snr loss? Hear Jour. 2000; 53:46–53.

Killion M, Niquette P, Gudmundsen G, Revit L, Banerjee S. Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners. J Acoust Soc Am. 2004; 116:2935–2405.

Kraus N, Chandrasekaran B. Music training for the development of auditory skills. Nat Rev Neurosci. 2010; 11:599–605. [PubMed: 20648064]

Krishnan A, Swaminathan J, Gandour JT. Experience-dependent enhancement of linguistic pitch representation in the brainstem is not specific to a speech context. Journal of Cognitive Neuroscience. 2008; 21:1092–1105. [PubMed: 18702588]

Krishnan A, Xu Y, Gandour J, Cariani P. Encoding of pitch in the human brainstem is sensitive to language experience. J Cogn Neurosci. 2005; 25:161–168.

Krizman J, Marian V, Shook A, Skoe E, Kraus N. Subcortical encoding of sound is enhanced in bilinguals and relates to executive function advantages. Proc Natl Acad Sci USA. 2012; 109:7877–7881. [PubMed: 22547804]

Landau SM, Marks SM, Mormino EC, Et Al. Association of lifetime cognitive engagement and low β-amyloid deposition. Arch Neurol. 2012; 69:623–629. [PubMed: 22271235]

Lunner T. Cognitive function in relation to hearing aid use. Int J Audiol. 2003; 42:S49–S58. [PubMed: 12918610]

Luo Y, Waite LJ. The impact of childhood and adult ses on physical, mental, and cognitive well-being in later life. J Gerontol B Psychol Sci Soc Sci. 2005; 60:S93–S101. [PubMed: 15746030]

Ma X, Suga N. Corticofugal modulation of the paradoxical latency shifts of inferior collicular neurons. J Neurophysiol. 2008; 100:1127–1134. [PubMed: 18596179]

Maguire EA, Gadian DG, Johnsrude IS, Good CD, Ashburner J, Frackowiak RS, Frith CD. Navigation-related structural change in the hippocampi of taxi drivers. Proc Natl Acad Sci USA. 2000; 97:4398–4403. [PubMed: 10716738]

Majno M. From the model of el sistema in venezuela to current applications: Learning and integration through collective music education. Ann N Y Acad Sci. 2012; 1252:56–64. [PubMed: 22524340]

Matilainen LE, Talvitie SS, Pekkonen E, Alku P, May PJC, Tiitinen H. The effects of healthy aging on auditory processing in humans as indexed by transient brain responses. Clin Neurophysiol. 2010; 121:902–911. [PubMed: 20359943]

Mazelová J, Popelar J, Syka J. Auditory function in presbycusis: Peripheral vs. Central changes Exp Gerontol. 2003; 38:87–94.

Mccoy SL, Tun PA, Cox LC, Colangelo M, Stewart RA, Wingfield A. Hearing loss and perceptual effort: Downstream effects on older adults' memory for speech. Q J Exp Psychol A. 2005; 58:22–33. [PubMed: 15881289]

Mesgarani N, Chang EF. Selective cortical representation of attended speaker in multi-talker speech perception. Nature. 2012; 485:233–236. [PubMed: 22522927]

Mukari S, Mamat WHW. Medial olivocochlear functioning and speech perception in noise in older adults. Audiol Neurootol. 2008; 13:328–334. [PubMed: 18460868]

Nadel L, Moscovitch M. Memory consolidation, retrograde amnesia and the hippocampal complex. Curr Opin Neurobiol. 1997; 7:217–227. [PubMed: 9142752]

Nahum M, Nelken I, Ahissar M. Low-level information and high-level perception: The case of speech in noise. PLoS Biol. 2008; 6:e126. [PubMed: 18494561]

Netz Y, Dwolatzky T, Zinker Y, Argov E, Agmon R. Aerobic fitness and multidomain cognitive function in advanced age. Int Psychogeriatr. 2011; 23:114–124. [PubMed: 20566000]

Nilsson M, Soli S, Sullivan J. Development of the hearing in noise test for the measurement of speech reception thresholds in quiet and in noise. J Acoust Soc Am. 1994; 95:1085–1099. [PubMed: 8132902]

Osler M, Avlund K, Mortensen EL. Socio-economic position early in life, cognitive development and cognitive change from young adulthood to middle age. Eur J Public Health. 2012

Parbery-Clark A, Anderson S, Hittner E, Kraus N. Musical experience offsets age-related delays in neural timing. Neurobiol Aging. 201210.1016/j.neurobiolaging.2011.12.015

Parbery-Clark A, Skoe E, Kraus N. Musical experience limits the degradative effects of background noise on the neural processing of sound. J Neurosci. 2009; 29:14100–14107. [PubMed: 19906958]

Parbery-Clark A, Strait DL, Anderson S, Hittner E, Kraus N. Musical experience and the aging auditory system: Implications for cognitive abilities and hearing speech in noise. Plos ONE. 2011; 6:e18082. [PubMed: 21589653]

Parbery-Clark A, Marmel F, Bair J, Kraus N. What subcortical–cortical relationships tell us about processing speech in noise. Eur J Neurosci. 2011; 33:549–557. [PubMed: 21255123]

Parthasarathy A, Bartlett EL. Age-related auditory deficits in temporal processing in f-344 rats. Neurosci. 2011; 192:619–630.

Patel AD. Why would musical training benefit the neural encoding of speech? The opera hypothesis. Front Psychol. 2011; 2:142. [PubMed: 21747773]

Peelle JE, Troiani V, Grossman M, Wingfield A. Hearing loss in older adults affects neural systems supporting speech comprehension. J Neurosci. 2011; 31:12638–12643. [PubMed: 21880924]

Pichora-Fuller K. Use of supportive context by younger and older adult listeners: Balancing bottom-up and top-down information processing. Int J Audiol. 2008; 47:S72–S82. [PubMed: 19012114]

Pichora-Fuller M, Schneider B, Daneman M. How young and old adults listen to and remember speech in noise. J Acoust Soc Am. 1995; 97:593–608. [PubMed: 7860836]

Pichora-Fuller MK. Cognitive aging and auditory information processing. Int J Audiol. 2003; 42:26–32.

Pichora-Fuller MK, Schneider BA, Benson NJ, Hamstra SJ, Storzer E. Effect of age on detection of gaps in speech and nonspeech markers varying in duration and spectral symmetry. J Acoust Soc Am. 2006; 119:1143–1155. [PubMed: 16521775]

Pichora-Fuller MK, Souza PE. Effects of aging on auditory processing of speech. Int J Audiol. 2003; 42:11–16.

Poth EA, Boettcher FA, Mills JH, Dubno JR. Auditory brainstem responses in younger and older adults for broadband noises separated by a silent gap. Hear Res. 2001; 161:81–86. [PubMed: 11744284]

Radua J, Phillips ML, Russell T, Lawrence N, Marshall N, Kalidindi S, El-Hage W, Mcdonald C, Giampietro V, Brammer MJ. Neural response to specific components of fearful faces in healthy and schizophrenic adults. Neuroimage. 2010; 49:939–946. [PubMed: 19699306]

Raizada RDS, Poldrack RA. Challenge-driven attention: Interacting frontal and brainstem systems. Front Hum Neurosci. 2007:2. [PubMed: 18958216]

Raykov T, Tomer A, Nesselroade JR. Reporting structural equation modeling results in psychology and aging: Some proposed guidelines. Psychol Aging. 1991; 6:499. [PubMed: 1777136]

Recanzone GH, Engle JR, Juarez-Salinas DL. Spatial and temporal processing of single auditory cortical neurons and populations of neurons in the macaque monkey. Hear Res. 2011; 271:115–122. [PubMed: 20430079]

Revelle, W. Psych: Procedures for personality and psychological research [Online]. Northwestern University; Evanston, IL: 2012. Available: http://personality-project.org/r/psych.manual.pdf

Revelle W, Zinbarg RE. Coefficients alpha, beta, omega, and the glb: Comments on sijtsma. Psychometrika. 2009; 74:145–154.

Rogers CS, Jacoby LL, Sommers MS. Frequent false hearing by older adults: The role of age differences in metacognition. Psychol Aging. 2012; 27:33–45. [PubMed: 22149253]

Rosseel Y. Lavaan: An r package for structural equation modeling. J Stat Softw. 2012; 48:1–36.

Rudner M, Rönnberg J, Lunner T. Working memory supports listening in noise for persons with hearing impairment. J Am Acad Audiol. 2011; 22:156–167. [PubMed: 21545768]

Russo N, Hornickel J, Nicol T, Zecker S, Kraus N. Biological changes in auditory function following training in children with autism spectrum disorders. Beh Brain Funct. 2010; 6:60.

Salthouse TA. Mental exercise and mental aging: Evaluating the validity of the "use it or lose it" hypothesis. Perspect Psychol Sci. 2006; 1:68–87.

Schellenberg EG. Music and cognitive abilities. Curr Dir Psychol Sci. 2005; 14:317–320.

Schellenberg EG, Peretz I. Music, language and cognition: Unresolved issues. Trends Cogn Sci. 2008; 12:45–46. [PubMed: 18178126]

Schvartz KC, Chatterjee M, Gordon-Salant S. Recognition of spectrally degraded phonemes by younger, middle-aged, and older normal-hearing listeners. J Acoust Soc Am. 2008; 124:3972. [PubMed: 19206821]

Shinn-Cunningham BG, Best V. Selective attention in normal and impaired hearing. Trends Amplif. 2008; 12:283–299. [PubMed: 18974202]

Shtyrov Y, Kujala T, Ahveninen J, Tervaniemi M, Alku P, Ilmoniemi RJ, Näätänen R. Background acoustic noise and the hemispheric lateralization of speech processing in the human brain: Magnetic mismatch negativity study. Neurosci Lett. 1998; 251:141–144. [PubMed: 9718994]

Skoe E, Kraus N. A little goes a long way: How the adult brain is shaped by musical training in childhood. J Neurosci. 2012; 32:11507–11510. [PubMed: 22915097]

Smith GE, Housen P, Yaffe K, Ruff R, Kennison RF, Mahncke HW, Zelinski EM. A cognitive training program based on principles of brain plasticity: Results from the improvement in memory with plasticity-based adaptive cognitive training (impact) study. J Am Ger Soc. 2009; 57:594–603.

Song JH, Skoe E, Banai K, Kraus N. Training to improve hearing speech in boise: Biological mechanisms. Cereb Cortex. 2012; 22:1180–1190. [PubMed: 21799207]

Souza P, Boike K, Witherell K, Tremblay K. Prediction of speech recognition from audibility in older listeners with hearing loss: Effects of age, amplification, and background noise. J Am Acad Audiol. 2007; 18:54–65. [PubMed: 17252958]

Streiner DL. Starting at the beginning: An introduction to coefficient alpha and internal consistency. J Pers Assess. 2003; 80:99–103. [PubMed: 12584072]

Swaab D. Brain aging and alzheimer's disease, "wear and tear" versus "use it or lose it". Neurobiol Aging. 1991; 12:317–324. [PubMed: 1755879]

Tierney A, Bergeson T, Pisoni D. Effects of early musical experience on auditory sequence memory. Empir Musicol Rev. 2008; 3:178–186. [PubMed: 21394231]

Toscano JC, Mcmurray B, Dennhardt J, Luck SJ. Continuous perception and graded categorization. Psychol Sci. 2010; 21:1532–1540. [PubMed: 20935168]

Tremblay K, Piskosz M, Souza P. Aging alters the neural representation of speech cues. Neuroreport. 2002; 13:1865–1870. [PubMed: 12395081]

Valenzuela MJ, Matthews FE, Brayne C, Ince P, Halliday G, Kril JJ, Dalton MA, Richardson K, Forster G, Sachdev PS. Multiple biological pathways link cognitive lifestyle to protection from dementia. Biol Psychiatry. 2011

Van Engen K, Bradlow A. Sentence recognition in native- and foreign-language multi-talker background noise. J Acoust Soc Am. 2007; 121:519–526. [PubMed: 17297805]

Vander Werff KR, Burns KS. Brain stem responses to speech in younger and older adults. Ear Hear. 2011; 32:168–180. [PubMed: 21052004]

Wagner W, Frey K, Heppelmann G, Plontke SK, Zenner HP. Speech-in-noise intelligibility does not correlate with efferent olivocochlear reflex in humans with normal hearing. Acta Otolaryngol. 2008; 128:53–60. [PubMed: 17851961]

Wakin, DJ. The New York Times. 2012 Feb 15. Fighting poverty, armed with violins.

Walton JP, Frisina RD, O'neill WE. Age-related alteration in processing of temporal sound features in the auditory midbrain of the cba mouse. J Neurosci. 1998; 18:2764–2776. [PubMed: 9502833]

Wilson RH, Mcardle RA, Smith SL. An evaluation of the bkb-sin, hint, quicksin, and win materials on listeners with normal hearing and listeners with hearing loss. J Speech Lang Hear Res. 2007; 50:844–856. [PubMed: 17675590]

Wingfield A, Tun PA. Cognitive supports and cognitive constraints on comprehension of spoken language. J Am Acad Audiol. 2007; 18:548–558. [PubMed: 18236643]

Wingfield A, Tun PA, Koh CK, Rosen MJ. Regaining lost time: Adult aging and the effect of time restoration on recall of time-compressed speech. Psychol Aging. 1999; 14:380–389. [PubMed: 10509694]

Wingfield A, Tun PA, Mccoy SL. Hearing loss in older adulthood. Current Directions in Psychological Science. 2005; 14:144–148.

Wong PCM, Ettlinger M, Sheppard JP, Gunasekera GM, Dhar S. Neuroanatomical characteristics and speech perception in noise in older adults. Ear Hear. 2010; 31:471–479. [PubMed: 20588117]

Wong PCM, Jin JX, Gunasekera GM, Abel R, Lee ER, Dhar S. Aging and cortical mechanisms of speech perception in noise. Neuropsychologia. 2009; 47:693–703. [PubMed: 19124032]

Wong PCM, Skoe E, Russo NM, Dees T, Kraus N. Musical experience shapes human brainstem encoding of linguistic pitch patterns. Nat Neurosci. 2007; 10:420–422. [PubMed: 17351633]

Woodcock, RW.; Mcgrew, KS.; Mather, N. Woodcock-johnson iii tests of cognitive abilities. Itasca, IL: Riverside Publishing; 2001.

Zatorre RJ, Belin P. Spectral and temporal processing in human auditory cortex. Cereb Cortex. 2001; 11:946–953. [PubMed: 11549617]

Zendel BR, Alain C. Musicians experience less age-related decline in central auditory processing. Psychol Aging. 2012; 27:410–417. [PubMed: 21910546]

Zhu, J.; Garcia, E. The wechsler abbreviated scale of intelligence (wasi). New York: Psychological Corporation; 1999.

Zinbarg R, Li W. Cronbach's alpha, revelle's beta, and mcdonald's omega: Their relations with each other and two alternative conceptualizations of reliability. Psychometrika. 2005; 70:123–133.

Highlights

- Cognition plays a strong role in a model of speech-in-noise perception.

- Central processing contributes to the older individual's ability to hear in noise.

- Central processing is mediated by cognitive function and life experiences.

- Early musical training leads to the use of different speech processing algorithms.
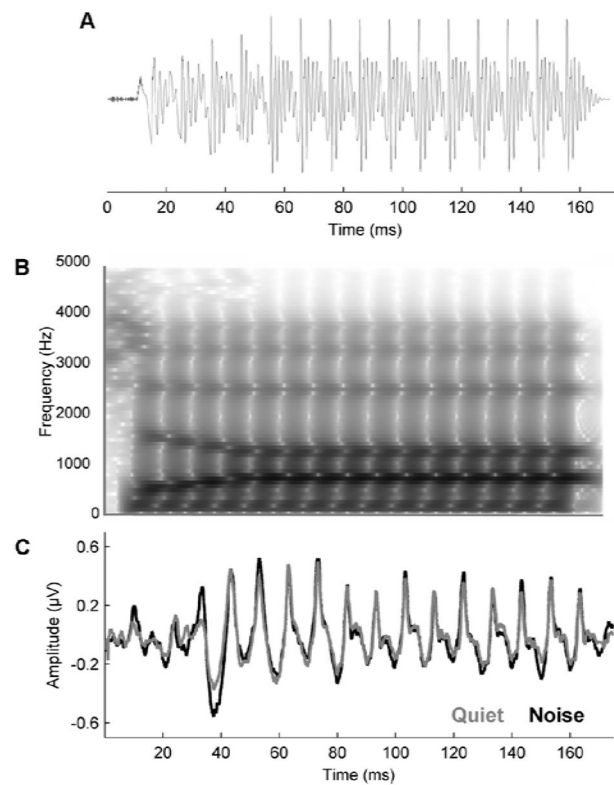
Figure 1.
The [da] stimulus time-amplitude waveform (A), spectrum (B) and the grand average cABR waveform (N = 120; C) obtained to the syllable presented in quiet (gray) and in six-talker babble (black).
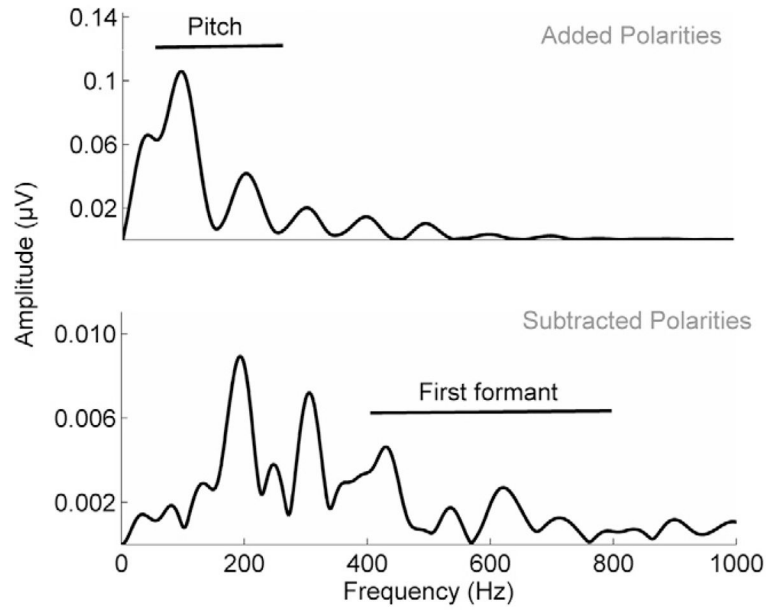
Figure 2.
Spectra transforms calculated for 20 to 60 ms of the response to the stimulus envelope (obtained by adding alternating polarities) (top) and temporal fine structure (obtained by subtracting alternating polarities) (bottom) of the stimulus [da] in 60 Hz bins. The lines indicate the frequencies over which the Pitch Encoding and F_1 Encoding variables were computed.
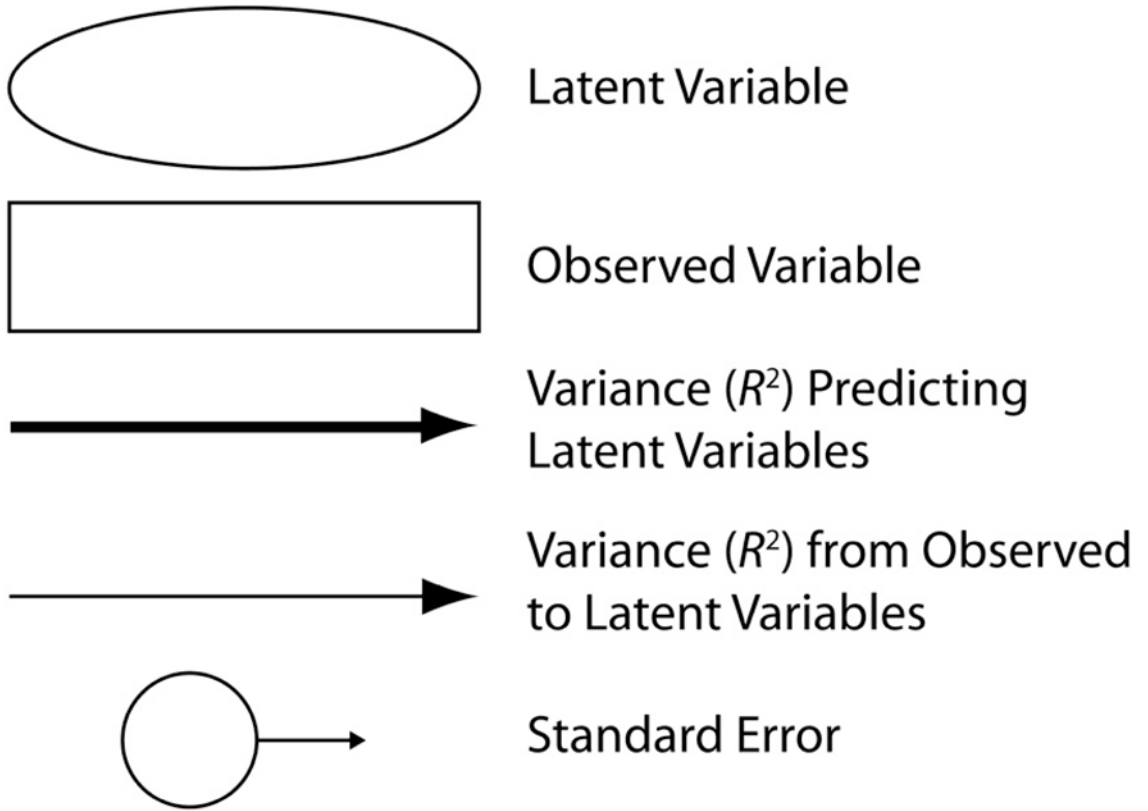
Figure 3.
The symbols and lines in Figures 4–6 are defined.
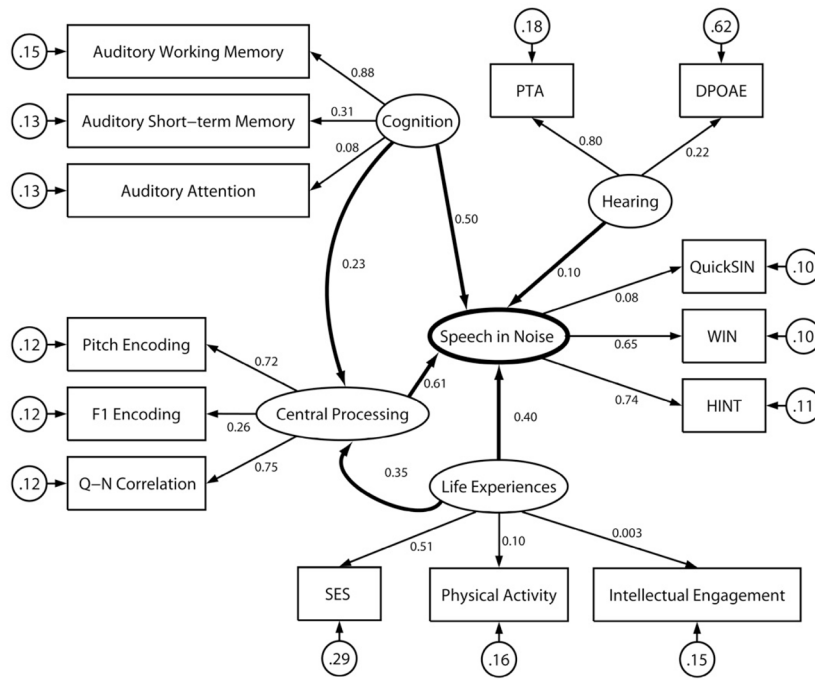
Figure 4.
In our structural equation model of Speech in Noise, Cognition and Central processing predict the greatest proportion of variance in the model, with Cognition and Life experiences having additional indirect effects on Speech-in-Noise performance through modulation of Central Processing. The larger R^2 values indicate stronger contributions from observed to latent variables and from latent variables to Central Processing or to Speech in Noise. The model accounted for 56.7% of the variance in Speech-in-Noise performance. Refer to Figure 3 for definitions of symbols and lines.
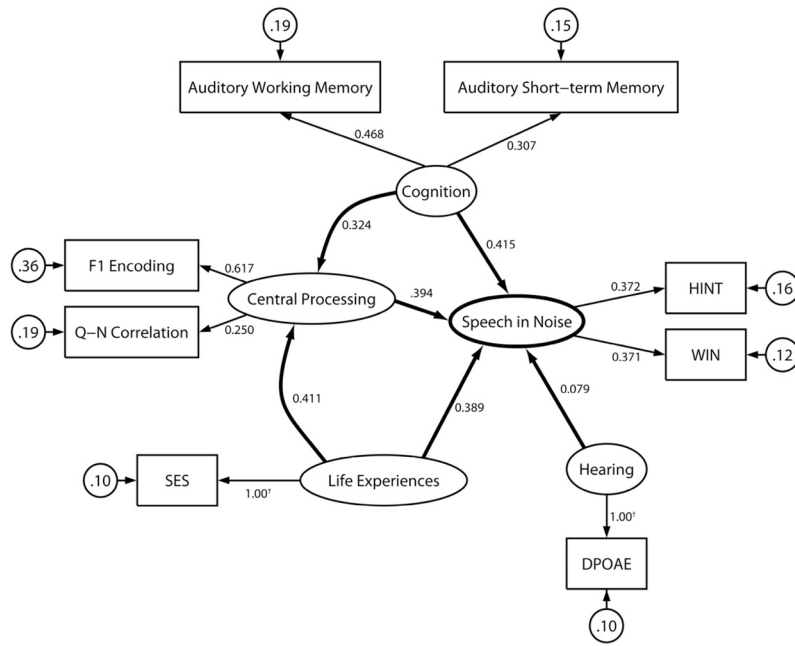
Figure 5.
Model 2 is cross-validated with a random sample of 60% of the participants (N=72). Cognition and Central Processing continue to contribute significantly to the variance; Life Experiences contribute to a greater extent than originally estimated in Model 2. This model accounted for 36.4% of the variance in Speech in Noise. Refer to Figure 3 for definitions of symbols and lines.
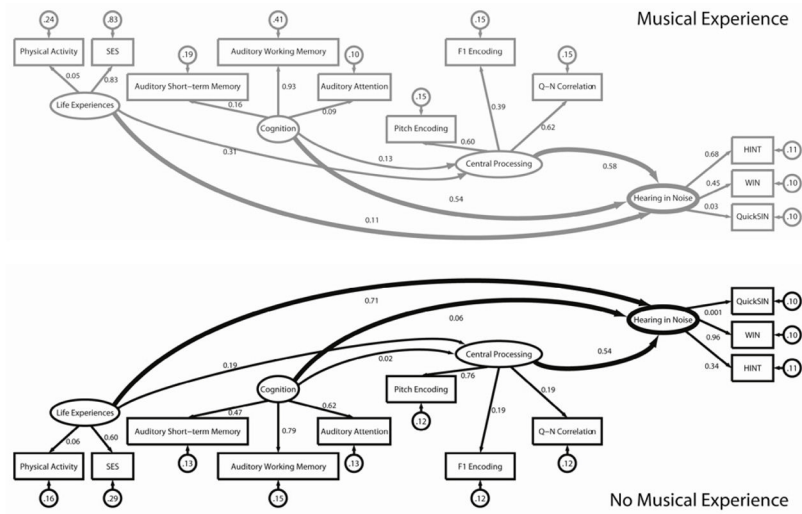
Figure 6.
Compensatory mechanisms for understanding speech in noise differ depending on life experiences: the Music+ group relies on cognition while the Music− group relies on life experience factors. Central processing plays an important role in both of the groups. Refer to Figure 3 for definitions of symbols and lines.

Table 1

Overall means (SD's) are displayed for each observed variable, as well as means (SD's) for the Music+ and Music− groups. Items excluded from the 2-group model (see 3.3) are presented in grey.

Item	Overall mean (S.D.)	Music+	Music−
QuickSIN (dB SNR loss)	1.31 (1.21)	1.24 (1.28)	1.37 (1.14)
WIN (dB SNR loss)	4.94 (2.23)	4.85 (2.36)	5.05 (2.11)
HINT (dB SNR)	−1.74 (1.38)	−1.74 (1.38)	−1.75 (1.41)
Auditory Attention (standard score)	103.4 (15.5)	103.02 (12.25)	104.49 (16.66)
Auditory Short-term Memory (standard score)	108.7 (12.7)	111.75 (13.80)	104.78 (10.11)
Auditory Working Memory (standard score)	113.0 (10.0)	114.40 (10.16)	111.27 (9.59)
Audiometry (PTA, .5–4 kHz, dB HL)	20.28 (9.36)	21.56 (10.00)	18.86 (8.35)
DPOAEs (dB SPL)	−9.26 (10.95)	−11.78 (11.35)	−6.55 (9.3)
Physical Activity	2.56 (1.30)	2.70 (1.27)	2.38 (1.33)
Intellectual Engagement	4.63 (2.58)	4.52 (2.74)	4.78 (2.42)
Socioeconomic Status (SES)	5.48 (1.14)	5.16 (0.14)	5.72 (0.16)
Pitch Encoding (μV)	0.22 (0.10)	0.24 (0.11)	0.20 (0.08)
F_1 Encoding (μV)	0.09 (0.04)	0.09 (0.04)	0.08 (0.04)
Q-N correlation (Pearson's r)	0.63 (0.18)	0.58 (0.16)	0.66 (0.18)

Table 2

Summary of latent variables, constinuent items, and reliability estimates. Factor-level reliability estimates are provided, namely: ω_h, ω_t, and a; see 2.9 for details on their estimation and interpretation.

Latent Variable	Item	ωh	ωt	a
Speech in Noise	QuickSIN	0.63	0.78	0.47
	WIN			
	HINT			
Cognition	Auditory Attention	0.62	0.81	0.57
	Auditory Short-term Memory			
	Auditory Working Memory			
Hearing	PTA	0.80	0.80	0.80
	DPOAEs			
Life Experiences	Physical Activity	0.41	0.60	0.22
	Intellectual Engagement			
	SES			
Central processing	Pitch Encoding	0.78	0.86	0.73
	F_1 Encoding			
	Q-N correlation			

Table 3

Item loadings for a five-factor model. A generalized least squared exploratory factor analysis with Kaiser normalization was estimated. Factor loadings are reported with corresponding latent factors estimated in subsequent structural models. Items assigned to each factor are presented in boldface.

	Speech in Noise	Cognition	Hearing	Life experiences	Central processing
QuickSIN	**−0.007**	−0.126	−0.062	0.074	0.138
WIN	**0.819**	−0.095	0.042	−0.003	0.197
HINT	**0.668**	−0.288	0.062	−0.042	0.135
Auditory Attention	0.138	**−0.217**	0.209	0.073	−0.042
Auditory Short-term Memory	−0.083	**0.664**	0.053	−0.095	−0.026
Auditory Working Memory	−0.235	**0.815**	−0.103	0.119	0.010
Audiometry	0.094	−0.053	**0.931**	0.137	−0.101
DPOAEs	0.024	−0.027	**−0.699**	0.068	0.021
Physical Activity	−0.130	0.078	0.032	**0.162**	0.110
Intellectual Engagement	−0.011	−0.022	0.056	**0.997**	0.032
SES	−0.245	0.058	0.100	**−0.099**	0.276
Pitch Encoding	0.077	−0.068	−0.129	−0.020	**0.822**
F1 Encoding	0.156	−0.126	0.004	0.146	**0.496**
Q-N Correlation	0.165	0.095	−0.028	0.029	**0.772**

Table 4

Relationships among Cognitive, Hearing, Speech-in-Noise, and Central Processing measures. The negative relationship between auditory attention and memory measures is due to a transformation we applied to correct the skewness of Auditory Attention scores.

	QuickSIN	WIN	HINT	Auditory Attention	Short-term Memory	Auditory Working Memory	PTA	DPOAEs	Physical Activity	Intellectual Engagement	SES	Pitch Encoding	F₁ Encoding	Q-N Correlation
QuickSIN	1	0.218*	0.162	-0.033	-0.418	-0.288**	0.118	0.054	0.062	-0.039	-0.076	-0.030	-0.017	0.008
WIN	0.218*	1	0.600***	0.103	-0.160	-0.271**	0.103	-0.005	-0.095	-0.001	-0.167	-0.232*	-0.242**	0.275**
HINT	0.162	0.600***	1	0.265**	-0.215*	-0.419***	0.112	-0.054	-0.060	-0.037	-0.067	0.154	0.227*	0.189*
Auditory Attention	-0.033	0.103	0.265**	1	-0.156	-0.215*	-0.272**	-0.118	0.057	0.096	0.012	0.086	0.188	-0.065
Short-term Memory	-0.418	-0.160	-0.215*	-0.156	1	0.536***	-0.004	-0.022	0.045	-0.107	0.051	-0.074	-0.166	0.044
Auditory Working Memory	-0.288**	-0.271**	-0.419***	-0.215*	0.536***	1	-0.156	0.050	0.112	0.098	0.108	-0.060	-0.073	0.037
PTA	0.118	0.103	0.112	-0.272**	-0.044	-0.156	1	-0.661***	0.007	0.190*	0.014	0.191*	0.004	-0.100
DPOAEs	0.054	-0.005	-0.054	-0.118	-0.022	0.050	-0.661***	1	-0.109	0.033	-0.120	0.120	0.122	0.020
Physical Activity	0.062	-0.095	-0.060	0.057	0.045	0.112	0.007	-0.109	1	0.169	0.176	0.057	0.042	0.066
Intellectual Engagement	-0.039	-0.001	-0.037	0.096	-0.107	0.098	0.190*	0.033	0.169	1	-0.083	0.001	0.165	0.049
SES	-0.076	-0.167	-0.067	0.012	0.051	0.108	0.014	-0.120	0.176	-0.083	1	0.1398	0.229*	0.154
Pitch Encoding	-0.030	-0.232*	0.154	0.086	-0.074	-0.060	0.191*	0.120	0.057	0.001	0.139	1	0.387***	0.654***
F₁ Encoding	-0.017	-0.242**	0.227*	0.188	-0.166	-0.073	0.004	0.122	0.042	0.165	0.229*	0.387***	1	0.372***
Q-N Correlation	0.008	0.275**	0.189*	-0.065	0.044	0.037	-0.100	0.020	0.066	0.049	0.154	0.654***	0.372***	1

*
p < 0.05;

**
p < 0.01;

***
p < 0.001

Table 5

Summary of Hierarchical Regression Analysis for Variables Predicting Speech-in-Noise Perception (HINT) (N=120)

Unstandardized (B and SE B) and standardized (β) coefficients in a model of contributions from observed variables contributing to the latent variables of Cognition, Life Experiences, and Central Processing. The overall model of speech-in-noise perception (based on the HINT score) is an appropriate fit, with Auditory Working Memory and Q-N Correlation providing significant contributions to the model.

Variable	B	SE B	β	p
Auditory Working Memory	−0.053	0.015	−0.357	0.001
Auditory Attention	0.159	0.106	0.152	0.138
F_1 Encoding	1.231	0.801	0.161	0.128
Q-N Correlation	0.245	0.566	0.053	0.043
Physical activity	−0.040	0.110	−0.035	0.716
SES	−0.135	0.135	−0.099	0.321

Table 6

Values from the cross-validation of Music+ and Music− models that include either Life Experiences or Cognition variables in the prediction of speech-in-noise performance. The variables in gray are latent variables rather than observed variables. The Music+_Cog and Music−+Life models (values in bold) are better predictors of speech-in-noise performance and are better models of the data based on χ^2 testing.

	Music+_Cog		Music+_Life		Music−_Cog		Music−_Life	
	R^2	SE variance	R^2	SE variance	R^2	SE variance	R^2	SE variance
WIN	0.389	0.154	0.653	0.373	0.964	0.178	0.868	0.132
HINT	0.873	0.206	0.486	0.290	0.615	0.142	0.656	0.124
QuickSIN	0.005	0.203	0.980	0.192	0.051	0.213	0.028	0.218
Auditory Working Memory	0.952	0.293			0.538	0.169		
Short-term Memory	0.323	0.170			0.611	0.173		
Auditory Attention	0.023	0.200			0.389	0.170		
Physical Activity			0.914	0.194			0.180	0.207
SES			0.773	0.274			0.377	0.244
Pitch	0.741	0.145	0.311	0.135	0.811	0.146	0.720	0.155
F_1 Formant	0.241	0.164	0.693	0.150	0.258	0.175	0.308	0.171
Q-N Correlation	0.732	0.145	0.359	0132	0.653	0.138	0.616	0.149
Speech in Noise	0.443		0.109		0.353		0.793	
Central Processing	0.009		0.721		0.010		0.008	
$\div^2$ (chi-square) degrees of freedom	26.64 (24)		11.36 (17)		22.36 (24)		9.49 (17)	

Table 7

Summary of "Enter" Hierarchical Regression Analysis for Variables Predicting Speech-in-Noise Perception in Music− (N=52) and Music+ (N=68) groups.

Standardized (β) coefficients in models comparing separate contributions ($3R^2$) from the SES variable on the first step and Auditory Working Memory, Q-N Correlation, and Physical Activity on the second step. For both the Music+ and Music− groups, after controlling for SES, the most important factor was Auditory Working Memory.

Music+			
Variable	**ΔR²**	**β**	**p**
Model 1			
SES	0.001	0.035	0.794
	ΔR²	β	p
Model 2	0.316		< 0.001
Auditory Working Memory		0.521	< 0.001
Q-N correlation		0.184	0.131
Physical Activity		0.170	0.150
Music−			
Variable	3R²	β	p
Model 1			
SES	0.071	0.266	0.074
	ΔR²	β	p
Model 2	0.161		0.048
Auditory Working Memory		0.311	0.034
Q-N Correlation		0.147	0.293
Physical Activity		0.168	0.258