

# Cognitive predictors of perceptual adaptation to accented speech

Briony Banks<sup>a)</sup>

*School of Psychological Sciences, University of Manchester, Manchester M13 9PL, United Kingdom*

Emma Gowen

*Faculty of Life Sciences, University of Manchester, Manchester M13 9PL, United Kingdom*

Kevin J. Munro and Patti Adank<sup>b)</sup>

*School of Psychological Sciences, University of Manchester, Manchester M13 9PL, United Kingdom*

(Received 8 July 2014; revised 19 February 2015; accepted 2 March 2015)

The present study investigated the effects of inhibition, vocabulary knowledge, and working memory on perceptual adaptation to accented speech. One hundred young, normal-hearing adults listened to sentences spoken in a constructed, unfamiliar accent presented in speech-shaped background noise. Speech Reception Thresholds (SRTs) corresponding to 50% speech recognition accuracy provided a measurement of adaptation to the accented speech. Stroop, vocabulary knowledge, and working memory tests were performed to measure cognitive ability. Participants adapted to the unfamiliar accent as revealed by a decrease in SRTs over time. Better inhibition (lower Stroop scores) predicted greater and faster adaptation to the unfamiliar accent. Vocabulary knowledge predicted better recognition of the unfamiliar accent, while working memory had a smaller, indirect effect on speech recognition mediated by vocabulary score. Results support a top-down model for successful adaptation to, and recognition of, accented speech; they add to recent theories that allocate a prominent role for executive function to effective speech comprehension in adverse listening conditions. © 2015 Acoustical Society of America. [<http://dx.doi.org/10.1121/1.4916265>]

[CGC]

Pages: 2015–2024

## I. INTRODUCTION

The ability to recognize speech in adverse listening conditions is a robust and flexible mechanism that is supported by our ability to “tune in” to unfamiliar or distorted speech (for reviews, see Samuel and Kraljic, 2009; Cristia *et al.*, 2012; Mattys *et al.*, 2012). Such perceptual adaptation can be defined as improved speech recognition (that is, accessing the semantic content of the speech message through perceiving the acoustic signal) as a result of exposure to an unfamiliar speech type. Despite the robustness of this ability, the relative success of perceptual adaptation can vary, and may depend on individual differences in the cognitive ability of the listener.

While it is increasingly acknowledged that certain cognitive abilities (such as working memory or executive function) play an important role in perceptual adaptation to unfamiliar speech (Adank and Janse, 2010; Erb *et al.*, 2012; Huyck and Johnsrude, 2012; Janse and Adank, 2012), no comprehensive model exists to explain the cognitive mechanisms underlying this ability. Given that adapting to adverse listening conditions is an inherent part of human communication, understanding the mechanisms underlying perceptual adaptation will contribute to existing models of speech recognition as well as a growing body of research into

communication in adverse conditions, which is relevant to both healthy and clinical populations.

The role of cognition has been widely investigated in relation to auditory processing in normal-hearing and hearing-impaired populations (e.g., Pichora-Fuller and Singh, 2006), particularly for recognition of speech in noise (for a review, see Akeroyd, 2008). However, it is not known whether such findings translate to perceptual adaptation to unfamiliar speech, particularly in a young, normal-hearing population. Existing accounts of speech perception currently emphasize the role of working memory in optimal and adverse listening conditions; for example, the ease of language understanding model (Ronnberg *et al.*, 2008) proposes that in difficult conditions, memory *storage* is required to keep track of the unfolding speech signal, while memory *processing* is required when speech input does not match existing phonological representations. Although working memory is a relatively reliable predictor for recognition of speech-in-noise (for normal-hearing and hearing-impaired adults; Akeroyd, 2008), evidence for a strong relationship between working memory and adaptation to unfamiliar speech is limited. Janse and Adank (2012) observed a relationship between working memory and recognition of a novel accent; however, this has not been replicated for perception of non-native (Gordon-Salant *et al.*, 2013), frequency compressed (Ellis and Munro, 2013) or noise-vocoded (Erb *et al.*, 2012) speech. There are three possible explanations for this limited evidence. First, it could be that working memory does not play as prominent a role in perceptual adaptation to unfamiliar speech as predicted by the ease of language understanding model; indeed, the model

<sup>a)</sup>Author to whom correspondence should be addressed. Electronic mail: briony.banks@manchester.ac.uk

<sup>b)</sup>Current address: Division of Psychology and Language Sciences, University College London, Chandler House, 2 Wakefield Street, London WC1N 1PF, United Kingdom.

endeavors to predict ease of understanding rather than speech recognition *per se* (Ronnberg, 2003). Second, the effect of working memory may be relatively subtle and the aforementioned studies may not have had the required statistical power to detect a small effect. Third, perceptual adaptation to unfamiliar speech may be primarily driven by other cognitive abilities (such as executive function or linguistic abilities) while working memory may have a more indirect influence similar to that observed for speech reading (Lyxell and Ronnberg, 1989), or for perceptual adaptation to degraded visual input (Kennedy *et al.*, 2009).

Behavioral and neuroimaging research has indeed provided support for a role of executive function during perceptual adaptation to unfamiliar speech. Executive function has been defined as cognitive processes, such as inhibitory mechanisms, that help control and coordinate other aspects of cognition, and is associated with activity in the frontal lobe (e.g., Miyake *et al.*, 2000). Neuroimaging studies have revealed activity in cortical regions associated with executive function when processing degraded compared with clear speech (Wild *et al.*, 2012; Erb *et al.*, 2013), while behavioral studies have demonstrated that attentional mechanisms are recruited for perceptual adaptation in lower level auditory training (Halliday *et al.*, 2011), and higher level adaptation to noise-vocoded (Huyck and Johnsrude, 2012), frequency-compressed (Ellis and Munro, 2013), and accented speech (Adank and Janse, 2010; Janse and Adank, 2012). However, it is unclear exactly how executive functions contribute to perceptual adaptation. Attentional control may certainly aid the listener to direct attention to the more salient aspects of the perceived speech (Amitay, 2009), or to better attend to the cognitively demanding input. Nevertheless, this does not explain how perceivers are able to learn and adapt to the new speech patterns of an unfamiliar accent, particularly how perceptual ambiguities are resolved or how correct lexical items are identified and selected. Successful perceptual adaptation may therefore be supported by inhibitory processes that facilitate the identification of correct lexical items and inhibit incorrect responses. Although measures of inhibition have predicted successful speech recognition in noise (Sommers and Danielson, 1999; Janse, 2012; Koelewijn *et al.*, 2012), they have thus far not been related to perceptual adaptation to unfamiliar speech.

Linguistic abilities, and particularly processing of lexical information, may also contribute to perceptual adaptation to unfamiliar speech. Studies have demonstrated that the lexical positioning of ambiguous phonemes affects subsequent perceptual categorisation of that phoneme (Norris *et al.*, 2003; Eisner and McQueen, 2005) and that intact lexical information is important for adaptation to noise-vocoded speech (Davis *et al.*, 2005). Nevertheless, only one study to date has investigated individual vocabulary knowledge as a predictor of perceptual adaptation to unfamiliar speech; in a study of older adults, Janse and Adank (2012) observed that better vocabulary knowledge predicted greater adaptation to accented speech. Given that vocabulary knowledge is relatively preserved in an older population, particularly in comparison to working memory and executive function (Schaie *et al.*, 1994; Singer *et al.*, 2003), a reliance on vocabulary

knowledge in this population may reflect a compensatory strategy rather than the normal route to adaptation in younger adults. To confirm whether this finding generalizes to a wider population, it is therefore necessary to also test a younger, normal-hearing population as a baseline measure.

Given the evidence described above, we propose that inhibition and vocabulary knowledge substantially contribute to perceptual adaptation to unfamiliar speech, while working memory contributes to a lesser extent. These three abilities have not previously been tested together in a single model of perceptual adaptation, thus allowing for their relative individual importance, as well as their combined contribution, to be examined. Testing these abilities in a large sample from a young, healthy population will enable detection of smaller effects while controlling for confounding factors of age-related sensory and cognitive decline. Furthermore, previous research has either focused on overall recognition of unfamiliar speech, or on adaptation (improvement in recognition accuracy) over time; we propose that these measures may tap into different cognitive processes and that both should be included in studies of speech perception in adverse listening conditions. The present study therefore investigated the contribution of three cognitive abilities (inhibition, vocabulary knowledge, and working memory) in adaptation to, and recognition of, accented speech. We chose to investigate accented speech as it is a naturalistic variant that is pertinent to everyday communication and, although adaptation to other distortions (such as noise-vocoded speech) likely involve the same mechanisms, it is not known whether they can be directly compared. We tested younger adults to build on previous results from older adults while providing baseline evidence from a cognitively healthy and normal-hearing population. Our hypothesis was that better abilities in the three cognitive measures would lead to greater and more rapid adaptation and to better overall recognition accuracy of the accented speech, with inhibition and vocabulary knowledge accounting for a greater amount of variance than working memory.

## II. METHOD

### A. Participants

One hundred students (24 male; mean age, 20.4 years; standard deviation, 2.28; range 18–30 years) recruited from the University of Manchester, participated in the study (for a linear multiple regression analysis with four predictor variables, a sample size  $>95$  is required to detect an effect size of 0.15 [ $\alpha = 0.05$ ,  $1 - \beta = 0.85$ ], Faul *et al.*, 2009). All participants were native British English speakers with no history of neurological, psychiatric, speech, or language problems (self-declared). Participants' hearing was assessed using pure-tone audiometry at 0.5, 1, 2, and 4 kHz in each ear separately. Any participant with a hearing threshold level  $>20$  dB for more than one frequency in either ear was excluded from the study. We provided compensation of course credit or £7.50 for participation. The study was approved by The University of Manchester ethics committee, and all participants gave their written informed consent.

## B. Materials

Stimulus material consisted of 105 Institute of Electrical and Electronics Engineers (IEEE) Harvard sentences (IEEE, 1969), selected because of their low predictability and standardized structure and length. We transcribed 90 of the sentences into a novel accent (Maye *et al.*, 2008; Adank and Janse, 2010). We chose to use a novel accent as a naturalistic stimulus that avoids confounds from participant familiarity and allows for a matched-guise design (Lambert *et al.*, 1960); that is, we could create stimuli from the same speaker in a standard and novel accent. The accent was created by systematically changing the vowel sounds of a standard British English accent, using vowel sounds from a variety of English regional accents (e.g., Scottish, Irish and Northern English; see Table I for the full phonetic transcription). This was achieved through an iterative process where we maintained the length of the vowel sounds (long, short or diphthongs) so as not to affect stress patterns. Our aim was to create an accent that would be unfamiliar to all participants but also of relatively low intelligibility (in order to measure adaptation over time, we required an accent with low intelligibility to avoid ceiling effects in earlier trials); to this end, some vowels sounds were not modified at all (that is, they remained as standard British English vowels). When asked about the accent after the experiment, the majority of participants indicated that it “sounded a bit like” an existing regional English accent (e.g., Scottish or Irish) but could not identify it.

A 30-year-old male speaker with a Standard British English accent was trained in the novel accent to provide all accented stimuli for the experiment. Recordings were made in a sound-treated laboratory with a SM58 microphone (Shure Inc., Niles, IL). All recordings were manually checked by the experimenter for pronunciation accuracy and naturalness, and any that were not deemed suitable (e.g., due to mispronunciation) were excluded from the study. Ninety

novel accented sentences were divided into 6 lists of 15 sentences to be used as the testing stimuli. A further 15 sentences recorded by the same speaker in a Standard British English accent were selected to be the baseline “unaccented” sentences (see Sec. II C for details). All audio files were normalized by equating the root-mean-square amplitude, resampled at 22 kHz in mono (over both ears) and cropped at the nearest zero crossings at voice onset and offset, using Praat software (Boersma and Weenink, 2012).

## C. Procedure

Participants wore sound attenuating headphones (HD 25-SP II; Sennheiser electronic GmbH & Co. KG, Wedemark, Germany) for the duration of the experiment. The volume level was adjusted to a comfortable level by the experimenter for the first participant and then kept at the same level for all participants thereafter. Stimuli were presented using MATLAB software (R2010a, MathWorks, Natick, MA; see Sec. II E for full details). To familiarize participants with the procedure, and to gain a baseline measurement of recognition accuracy for native speech, participants first listened to the 15 unaccented sentences as practice trials, followed by the 90 accented sentences. Sentence lists were counterbalanced across the six testing blocks, each comprising 15 sentences, and were presented in a pseudo-random order per testing block and per participant. Each sentence was presented once to each participant to avoid training effects of particular items. Last, participants were tested on the three cognitive measures. The experiment was carried out in one session lasting approximately 60 min. As part of a wider study, participants also underwent training with additional versions (audiovisual, audio-only or visual-only) of the novel-accented stimuli between block 3 and block 4; however, no significant effects of training were observed,<sup>1</sup> and these results will not be discussed further in this paper.

## D. Speech recognition task

After presentation of each sentence, we instructed participants to repeat out loud as much or as little of the sentence as they could, in their normal voice and without imitating the accent. The experimenter scored participants’ responses immediately after each trial according to how many keywords out of a possible four were correctly repeated. These responses were logged using MATLAB to determine the signal-to-noise ratio (SNR) of the next trial (see Sec. II E for details). No feedback was given to participants. Keywords comprised either content or function words and, in line with previous studies of perceptual adaptation to unfamiliar speech (Dupoux and Green, 1997; Golomb *et al.*, 2007), were marked as correct despite incorrect suffixes (such as -s, -ed, -ing) or verb endings. If only part of a word (including compound words) was repeated it was counted as incorrect. If a participant repeated a word imitating the novel accent (that is, if their pronunciation deviated from their own accent to match the novel accent), this was also counted as incorrect, as we could not ascertain whether the participant had correctly identified the lexical item, or whether they had simply repeated the phonological pattern they had heard.

TABLE I. Phonetic description of the novel accent.

International Phonetic Alphabet	Example
ɪ → ε	sit → set
ε → ɪ	bet → bit
æ → ε	hat → het
ʌ → ʊ	cud → could
ɜ: → εə	girl → gairl
a: → ɔ:	dark → dork
ɒ → ɔ:	hot → hawt
ɔ:	door
u:	food
ʊ	good
ə	mother
i:	tree
εə → ɜ:	hair → her
əʊ → aʊ	vote → vovt
aʊ → u:	how → hoo
εɪ → aɪ	way → wye
aɪ → ɔɪ	my → moy
ɪə	hear
ɔɪ	joy



## E. Speech reception thresholds

Recognition accuracy during each testing block was measured by establishing participants' Speech Reception Thresholds (SRTs) in speech-shaped background noise, using an adaptive staircase procedure (Plomp and Mimpen, 1979). Measuring speech recognition in this way avoids ceiling effects associated with rapid perceptual adaptation to accented speech, and also controls for variation in individual baseline comprehension. Accuracy (number of correctly repeated keywords) was maintained at 50% by adjusting the SNR in pre-determined steps. Thus, as perceptual adaptation took place and correct responses increased, the SNR was decreased and the task became increasingly difficult (Baker and Rosen, 2001). The procedure was carried out using MATLAB software. The initial SNR for the first sentence in each block was 10 dB. Throughout the staircase procedure, the background noise varied in steps of 8 dB for the first two reversals, and 2 dB for each reversal thereafter. The mean SNR for all reversals per testing block indicated the SRT measurement for each participant.

## F. Cognitive background measures

Vocabulary knowledge was tested using the Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999) vocabulary subtest, which requires participants to provide oral definitions of words. Participants were scored according to the standard instructions, and overall percentages were calculated for analysis. Inhibition was measured using a standard Stroop test (Stroop, 1935), presented to the participant on paper and requiring oral responses. The test comprised three sections: Color naming (C), word naming (W), and word-color interference (WC), whereby participants were required to name the (incongruent) color of the ink that words were written in. Each section was timed manually by the experimenter using a stopwatch. Interference scores, based on the mean time (in seconds) to complete each section, were calculated using the following equation:

$$\text{Interference} = \text{WC} - (\text{W} \times \text{C}) / (\text{W} + \text{C}).$$

Finally, working memory was tested using an English version of a standard reading span test (Ronnberg *et al.*, 1989). This requires participants to read 3–6 sentences which appear on screen word-by-word, and then to subsequently recall either the first or last word of each sentence when prompted by the experimenter. The total number of correctly recalled words was calculated for analysis.

## G. Data analysis

Within our data set, we identified two outliers (one for the accented SRTs and one for the unaccented SRTs) with standardized residuals  $>3.29$ , and these scores were modified to the value of the group mean SRT plus two standard deviations. Interference scores for the Stroop test were positively skewed, so the data were log transformed to allow for parametrical analysis. Mauchly's test indicated that the assumption of sphericity had been violated for the repeated

measures analysis of variance (ANOVA),  $\chi^2(14) = 75.61$ ,  $p < 0.001$ , therefore degrees of freedom were corrected using Huynh-Feldt estimates of sphericity ( $\epsilon = 0.86$ ). Unless otherwise stated, all other assumptions for parametrical testing of the data were met.

Recognition of unfamiliar speech can be measured in two ways: As overall performance, or as improvement in performance over time, and both of these measures were used in our analyses of individual differences. Overall performance (recognition accuracy) was calculated as the mean SRT across all testing blocks. Adaptation was analyzed as the amount and rate of improvement. We calculated the amount of adaptation as the difference in mean SRTs between the first three and the last three testing blocks, while rate of adaptation was calculated by fitting a linear function to the recognition accuracy data (Erb *et al.*, 2012); we used the equation  $y = mx + b$ , where  $y$  is the mean SRT,  $x$  is time (block),  $m$  is the slope, and  $b$  is the intercept. The slope of each participant's linear fit was used as a measurement of adaptation rate. To investigate individual differences in perceptual adaptation, we used multiple linear regression to analyze the relationships between recognition accuracy, amount and rate of adaptation (our dependent variables), and four predictor variables: unaccented SRTs (representing participants' baseline ability to deal with speech in noise), and vocabulary, working memory and Stroop interference scores. We included unaccented SRTs in order to examine relationships between the cognitive predictors and comprehension when unaccented SRTs were held constant; that is, we could infer that the individual contribution of each cognitive measure was related to the accented speech over and above the background noise.

To test our hypothesis that working memory may have an indirect effect on comprehension (that is, that the relationship between working memory and comprehension was mediated by other predictors), we used path analysis, fitting a hypothesized model to our data and thus assessing the direct and indirect (mediated) effects between variables. Model fit was assessed using the chi-square ( $\chi^2$ ) statistic, the root-mean-square error of approximation (RMSEA), and the Tucker-Lewis Index (TLI). As our sample size was relatively small for this type of analysis, we used bootstrapping (Shrout and Bolger, 2002; Preacher and Hayes, 2004) to construct bias-corrected confidence intervals (95%) to test for mediation effects between variables.

## III. RESULTS

### A. Perceptual adaptation to accented speech

Table II shows the mean SRTs for the unaccented speech (hereafter "unaccented"), as well as SRTs for each testing block of the accented speech. As SRTs represent the signal-to-noise ratio (dB), higher levels reflect poorer tolerance to background noise (poorer performance). As expected, unaccented SRTs were significantly lower than mean accented SRTs for each testing block, even after correcting for multiple comparisons [block 1,  $t(99) = -21.20$ ,  $p < 0.001$ ; block 2,  $t(99) = -21.68$ ,  $p < 0.001$ ; block 3,  $t(99) = -14.76$ ,  $p < 0.001$ ; block 4,  $t(99) = -14.18$ ,  $p < 0.001$ ; block 5,

TABLE II. Mean SRTs and standard deviations per testing block.

Testing block	Mean (dB)	Standard deviation (dB)
Unaccented	0.57	1.7
1	10.18	4.54
2	7.54	3.16
3	5.48	3.50
4	6.54	4.42
5	5.60	2.95
6	4.03	2.71

$t(99) = -15.45$ ,  $p < 0.001$ ; block 6,  $t(99) = -13.14$ ,  $p < 0.001$ ], confirming that the novel accent negatively affected participants' performance. To confirm whether participants' tolerance to background noise significantly changed over time, we carried out a repeated-measures ANOVA to examine within-subject effects of testing block (6 levels). We observed a significant main effect of testing block [ $F(4.32, 409.88) = 45.72$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.33$ ]. Pairwise comparisons (Bonferroni correction,  $p < 0.003$ ) revealed that SRTs for blocks 2, 3, 4, 5 and 6 were significantly lower than SRTs in block 1, confirming that participants' tolerance to the background noise increased. SRTs for blocks 3, 5 and 6 were also significantly lower than block 2, and block 6 was significantly lower than blocks 3, 4 and 5. As expected, there was considerable individual variation between participants' SRTs throughout the experiment (see Fig. 1). There was a significant negative correlation between the slope and intercept of all linear fits ( $r = -0.79$ ,  $p < 0.001$ ), indicating that participants who initially performed worse improved the most.<sup>2</sup>

## B. Cognitive ability and perceptual adaptation to accented speech

Table III shows the correlation matrix between adaptation amount, adaptation rate and recognition accuracy for the accented speech, and the four predictor variables.

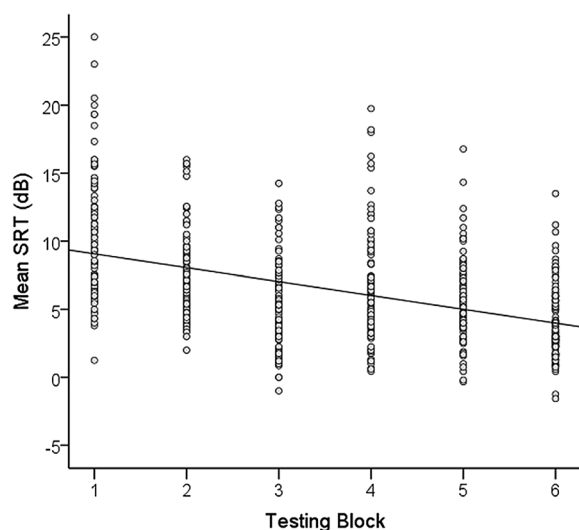


FIG. 1. Individual variation in recognition accuracy of accented speech in noise: Mean SRTs (in dB) per participant, per testing block, with mean linear fit for all participants.

Adaptation amount was negatively correlated with Stroop scores ( $r = -0.29$ ,  $p = 0.004$ ; see Fig. 2), indicating that lower interference scores (and thus better inhibition) was related to greater adaptation. Adaptation rate (slope) was positively correlated with Stroop scores ( $r = 0.21$ ,  $p = 0.04$ , indicating that better inhibition was related to a faster rate of adaptation (it should be noted that, as lower SRTs indicated better performance, adaptation slopes had mainly negative values ( $M = -1.01$ ); lower values of our adaptation rate measurement therefore represent faster adaptation). Recognition accuracy was positively correlated with unaccented SRTs ( $r = 0.36$ ,  $p < 0.001$ ), indicating that participants who could tolerate a high level of background noise for the unaccented sentences, could also tolerate a high level of background noise for the accented sentences. Recognition accuracy was negatively correlated with vocabulary ( $r = -0.38$ ,  $p < 0.001$ ) and working memory ( $r = -0.22$ ,  $p = 0.03$ ); that is, participants with better vocabulary and working memory scores had lower SRTs, and thus had better recognition accuracy of the accented speech. Between the four predictor variables, working memory was positively correlated with vocabulary (better working memory was related to greater vocabulary knowledge,  $r = 0.25$ ,  $p = 0.01$ ), and negatively correlated with Stroop interference scores (better working memory was related to greater inhibition,  $r = -0.25$ ,  $p = 0.01$ ). Vocabulary was negatively correlated with unaccented SRTs (greater vocabulary knowledge was related to better recognition accuracy of the unaccented sentences,  $r = -0.39$ ,  $p < 0.001$ ). Between the three outcome variables, recognition accuracy and adaptation rate were negatively correlated,  $r = -0.23$ ,  $p = 0.01$  (participants with poorer overall recognition accuracy adapted more quickly), and adaptation amount and rate were negatively correlated,  $r = -0.84$ ,  $p < 0.001$  (participants who adapted the most did so at a faster rate). No issues of collinearity were identified, and thus, all cognitive measures and the unaccented SRTs could be included in our regression analyses.

In order to analyze the contribution of the four predictor variables to recognition accuracy, adaptation amount and adaptation rate, we carried out three backward stepwise regression analyses. Table IV shows the results of the regression model for recognition accuracy of the accented speech. When all other predictor variables were held constant, unaccented SRTs ( $\beta = 0.27$ ,  $p = 0.008$ ) and vocabulary ( $\beta = -0.24$ ,  $p = 0.02$ ) significantly predicted recognition accuracy, whereas working memory did not ( $\beta = -0.16$ ,  $p = 0.09$ ). Table V shows the results of the regression models for adaptation amount and adaptation rate. In both models, Stroop scores (inhibition) significantly predicted the amount ( $\beta = -0.29$ ,  $p = 0.004$ ) and rate ( $\beta = 0.21$ ,  $p = 0.04$ ) of adaptation.

As we had observed a significant correlation between working memory and recognition accuracy, but working memory did not significantly predict recognition accuracy in our regression model, we hypothesized that there was an indirect relationship between these two variables, mediated by vocabulary score. We carried out a path analysis to test this hypothesis. The presence of correlations between the three variables (working memory, vocabulary and

TABLE III. Correlation matrix for recognition accuracy of, and adaptation to, accented speech and cognitive ability, with means and standard deviations ( $N = 100$ ).<sup>a</sup>

Variable	Mean	Standard deviation	Recognition	Adaptation amount	Adaptation rate	Unaccented	Vocabulary	Working memory	Stroop
Recognition Accuracy (SRT, dB)	6.50	2.05	–						
Adaptation amount (dB)	2.34	2.18	0.03	–					
Adaptation rate (slope)	–1.01	0.68	–0.23 <sup>b</sup>	–0.84 <sup>c</sup>	–				
Unaccented (SRT, dB)	0.57	1.70	0.36 <sup>c</sup>	0.03	–0.03	–			
Vocabulary (%)	66.76	6.88	–0.38 <sup>c</sup>	0.07	–0.04	–0.39 <sup>c</sup>	–		
Working Memory (%)	49.76	9.06	–0.22 <sup>b</sup>	0.13	–0.08	–0.01	0.25 <sup>b</sup>	–	
Stroop	1.52	0.09	0.15	–0.29 <sup>d</sup>	0.21 <sup>b</sup>	–0.09	–0.14	–0.25 <sup>b</sup>	–

<sup>a</sup>Higher mean scores for recognition accuracy and Stroop indicate poorer performance. Higher scores for all other variables indicate better performance.

<sup>b</sup>Two-tailed Pearson’s correlations, significant at  $p < 0.05$ .

<sup>c</sup>Two-tailed Pearson’s correlations, significant at  $p < 0.001$ .

<sup>d</sup>Two-tailed Pearson’s correlations, significant at  $p < 0.01$ .

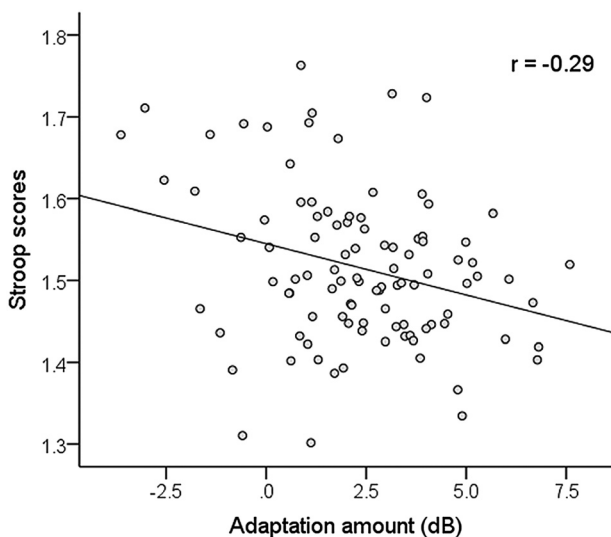


FIG. 2. Scatterplot showing correlation between amount of adaptation to accented speech and Stroop interference scores (inhibition), with linear regression best fit;  $r$  = correlation coefficient.

TABLE IV. Backward stepwise regression analysis for the predictors of recognition accuracy of accented speech ( $N = 100$ ).<sup>a</sup>

Variable	$B$	Standard error $B$	$\beta$
Step 1			
Unaccented SRTs	0.35	0.12	0.28 <sup>b</sup>
Vocabulary	–0.07	0.03	–0.22 <sup>c</sup>
Working memory	–0.03	0.02	–0.13
Stroop	2.50	2.04	0.12
Step 2			
Unaccented SRTs	0.33	0.12	0.27 <sup>b</sup>
Vocabulary	–0.07	0.03	–0.24 <sup>c</sup>
Working memory	–0.04	0.02	–0.16

<sup>a</sup> $R^2 = 0.24$  for Step 1;  $\Delta R^2 = -0.01$  for Step 2 ( $p < 0.05$ ).

<sup>b</sup> $p < 0.01$ .

<sup>c</sup> $p < 0.05$ .

recognition accuracy), meant that our data met the assumptions required for a mediation effect (Baron and Kenny, 1986). It should be noted that these assumptions were not met for the predictors of adaptation amount or rate, and so

TABLE V. Backward stepwise regression analysis for the predictors of (a) amount of adaptation and (b) rate of adaptation (slope) to accented speech ( $N = 100$ ).

Variable	$B$	Standard error $B$	$\beta$
V(a) Adaptation amount <sup>a</sup>			
Step 1			
Unaccented	0.02	0.14	0.01
Vocabulary	0.01	0.04	0.03
Working memory	0.01	0.03	0.05
Stroop	–6.22	2.36	–0.27
Step 2			
Vocabulary	0.01	0.03	0.02
Working memory	0.01	0.03	0.06
Stroop	–6.26	2.32	–0.27
Step 3			
Working memory	0.01	0.02	0.06
Stroop	–6.30	2.30	–0.27
Step 4			
Stroop	–6.64	2.22	–0.29 <sup>b</sup>
V(b) Adaptation rate <sup>c</sup>			
Step 1			
Unaccented	–0.01	0.05	–0.02
Vocabulary	0.00	0.01	0.01
Working memory	–0.01	0.01	–0.03
Stroop	1.40	0.76	0.19
Step 2			
Unaccented	–0.01	0.04	–0.01
Working memory	–0.01	0.01	–0.04
Stroop	1.40	0.75	0.20
Step 3			
Working memory	–0.01	0.01	–0.04
Stroop	1.41	0.74	0.20
Step 4			
Stroop	1.47	0.71	0.21 <sup>c</sup>

<sup>a</sup> $R^2 = 0.09$  for Step 1;  $\Delta R^2 = 0.00$  for Steps 2, 3 and 4 ( $p$ 's  $< 0.05$ ).

<sup>b</sup> $p < 0.05$ .

<sup>c</sup> $R^2 = 0.04$  for Step 1;  $\Delta R^2 = 0.00$  for Steps 2, 3 and 4 ( $p$ 's  $< 0.05$ ).

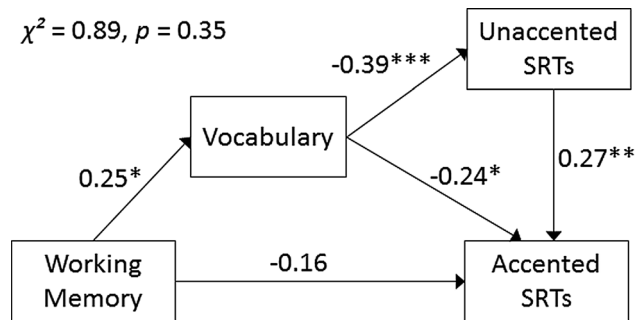


FIG. 3. Path analysis model for the cognitive predictors of recognition accuracy of accented speech. All path parameters are standardized coefficients (direct effects).  $\chi^2$  = chi-square statistic (non-significant value indicates the model is a good fit). The pathway between working memory and accented SRTs was not significant ( $p > 0.05$ ) and was mediated by vocabulary score. There was an indirect effect of working memory on accented SRTs,  $\beta = -0.09$ ,  $p < 0.01$ , and an indirect effect of vocabulary score on accented SRTs,  $\beta = -0.11$ ,  $p < 0.01$ . \*  $p < 0.05$ ; \*\*  $p < 0.01$ ; \*\*\*  $p < 0.001$ .

path analyses to test for mediation effects were not carried out on these data. Figure 3 shows the path model for the predictors of recognition accuracy with standardized coefficients. The inclusion of each pathway was based on observations from our data, while the direction of each pathway was based on our hypotheses (e.g., that vocabulary score predicted recognition accuracy). The model fit the data well:  $\chi^2(1) = 0.89$ ,  $p = 0.35$ ; TLI = 1.02; RMSEA < 0.001. As predicted, the relationship between working memory and recognition accuracy of the accented speech was mediated by vocabulary score; that is, working memory had an indirect effect on recognition accuracy,  $\beta = -0.09$ ,  $p < 0.01$ , via vocabulary score. Vocabulary had a direct effect on recognition accuracy,  $\beta = -0.24$ ,  $p < 0.01$ , and an indirect effect on recognition accuracy,  $\beta = -0.11$ ,  $p < 0.01$ , via unaccented SRTs; vocabulary therefore accounted for the greatest amount of total variance (combined direct and indirect effects) on recognition accuracy,  $\beta = -0.34$ ,  $p < 0.01$ .

#### IV. DISCUSSION

The present study investigated how individual differences in cognitive ability relate to perceptual adaptation to accented speech, as measured by overall performance (recognition accuracy) and amount of improvement (adaptation). We predicted that better inhibition (a measure of executive function) and vocabulary knowledge, supported by better working memory, would lead to better recognition accuracy and greater adaptation.

##### A. Perceptual adaptation to accented speech

As predicted from previous studies of adaptation to accented speech (Clarke and Garrett, 2004; Bradlow and Bent, 2008; Maye *et al.*, 2008; Adank and Janse, 2010; Gordon-Salant *et al.*, 2010; Janse and Adank, 2012), we observed significant improvements in recognition accuracy of our novel accent over time, represented by a greater tolerance to background noise in later compared to earlier trials. As expected, we observed considerable individual variation in SRTs throughout all testing blocks, and participants who had poorer starting levels adapted the most. Similar

adaptation patterns have been observed for comprehension of noise-vocoded speech (Stacey and Summerfield, 2007; Erb *et al.*, 2012).

Adaptation to accented speech can occur rapidly, even after as few as eight sentences (Clarke and Garrett, 2004). However, by using a relatively difficult novel accent and an adaptive procedure to vary the background and target SNR, this process was slowed; indeed, our participants continued to improve significantly until the final block of stimuli, after exposure to 90 sentences. The disadvantage of this procedure is that the measure of recognition accuracy obtained (SRTs) represents responses to the accented speech and to the background noise. Although we cannot completely separate both elements, several factors provide evidence that listeners adapted predominantly to the accent, and not to the background noise. First, mean SRTs for the accented speech were significantly different to SRTs for the unaccented speech; that is, participants never perceived the accented speech as well as the unaccented speech, even after exposure to all 90 test sentences. Second, Adank and Janse (2010) demonstrated that SRTs while listening to a standard native accent (using the same adaptive procedure as in the present study) remain stable in a young population, with a difference of < 1 dB in SRTs after exposure to 60 sentences. Third, neither of our adaptation measures was significantly correlated with unaccented SRTs, indicating that the amount and rate participants adapted was not related to their ability to process unaccented speech in background noise. This supports our claim that the adaptation we observed in our study (a mean improvement of 6 dB between the first and final testing blocks) was likely related to the accent rather than to the background noise. However, one further limitation should be acknowledged—that the perception of the same speaker with an unfamiliar accent, after listening to him speak with a standard British English accent, may have influenced the higher SRTs in the first block.

##### B. Cognitive ability and perceptual adaptation to accented speech

Our analyses revealed that inhibition, as measured by the Stroop test, predicted adaptation to the accented speech. Participants who had better inhibition (that is, performed better at the Stroop test) adapted more and at a faster rate than participants who demonstrated poorer inhibition, thus supporting our hypothesis. To our knowledge, ours is the first study to directly link inhibition to perceptual adaptation to accented speech. This finding adds to a growing body of evidence that executive function, such as inhibition or attention, has a major role in perceptual adaptation to unfamiliar speech (Huyck and Johnsrude, 2012; Wild *et al.*, 2012; Erb *et al.*, 2013), including adaptation to accented speech (Adank and Janse, 2010; Janse and Adank, 2012). Inhibitory abilities are likely recruited when competing (and incorrect) lexical responses are triggered by the accented speech (Brouwer *et al.*, 2012; Tuinman *et al.*, 2012), thus helping to resolve ambiguities in the speech signal. This may allow the listener to identify the correct lexical items and thus match unfamiliar phonemic patterns to existing phonemic



representations, resulting in adaption to the patterns of the accented speech. Greater inhibitory abilities may thus allow listeners to overcome ambiguous or unfamiliar auditory input such as accented speech.

Performance on the Stroop test has also been linked to recognition of speech in background noise in older adults (Sommers and Danielson, 1999; Janse, 2012). As our participants listened to the accented speech in background noise, this may explain part of the relationship between Stroop scores and adaptation observed in our study. However, if this were the case, we would also expect the Stroop scores and our adaptation measures to correlate with SRTs for the unaccented speech. No such correlations were observed, which indicates that the relationship between Stroop scores and adaptation reflects efficient adaptation to the accent rather than to the background noise. Nevertheless, it should be noted that our participants only listened to 15 unaccented sentences—fewer than in previous studies that observed a relationship between Stroop scores and speech recognition in noise (Sommers and Danielson, 1999; Janse, 2012); therefore, we may not have observed a correlation between unaccented SRTs and Stroop scores due to the small amount of exposure. A third possible interpretation of our findings is that the Stroop test relates to more than one aspect of executive function, or to individual strategies such as attention or motivation. Although it is not possible to separate the cognitive constructs of the Stroop test in this experiment, overall strategies such as motivation or attention would likely apply to all three cognitive predictors, whereas only Stroop scores were significantly related to adaptation.

Our second finding was that vocabulary knowledge predicted recognition accuracy of the accented speech. As we hypothesized, participants who had greater vocabulary scores could tolerate more background noise overall, and thus their recognition of the accented speech was more robust than participants with lower vocabulary scores. This confirms a role for vocabulary knowledge during perception of accented speech in a young, healthy population, and supports similar findings in older adults (Janse and Adank, 2012). Our path analysis revealed that vocabulary knowledge accounted for the greatest amount of total variance in recognition of the accented speech. We observed a direct relationship between vocabulary knowledge and recognition of the accented speech, but we also observed an indirect relationship via recognition of the unaccented speech (that is, unaccented SRTs partially mediated the relationship between vocabulary score and accented SRTs). Vocabulary score also fully mediated the relationship between working memory and recognition of the accented speech. This suggests a particular importance for lexical knowledge in successfully perceiving native and non-native speech in noise. Greater vocabulary knowledge likely allows the listener to more readily identify and access lexical items from unfamiliar or ambiguous auditory input; stronger mapping between lexical and semantic representations may also help listeners to process the incremental speech input by helping them to anticipate upcoming words in the sentence (Borovsky *et al.*, 2012). Although the role of lexical processing in perceptual adaptation to other speech distortions is debated, for

example, noise-vocoded (Hervais-Adelman *et al.*, 2008) and time-compressed (Janse, 2009) speech, lexical information may be particularly pertinent to comprehension of accented speech (e.g., Norris *et al.*, 2003), perhaps aiding the listener to identify patterns of phonetic variation by allowing them to map this variation more easily onto lexical items. However, a second interpretation of our finding is also possible. Vocabulary knowledge is usually correlated with verbal and non-verbal IQ (Wechsler, 1958; Kamphaus, 2005), and indeed, the test used in our study is part of a standard IQ test battery. Our findings here may thus reflect a relationship between speech recognition and general intelligence, rather than specifically with vocabulary knowledge, although measures of IQ have not consistently been found to predict recognition of native speech in noise (Akeroyd, 2008). As we did not test our participants' full IQ, further investigation is required to confirm whether lexical knowledge in particular, or general intelligence, are important for successful recognition of accented speech.

Vocabulary knowledge did not predict amount or rate of adaptation to the accented speech as we had hypothesized, which is contrary to results observed in older adults (Janse and Adank, 2012). These discrepant findings may reflect differences in the populations tested; as vocabulary knowledge can increase into the sixth decade (Schaie *et al.*, 1994) and remains relatively stable into the eighth (Singer *et al.*, 2003), it may provide an important compensatory strategy in older adults following a decline in other cognitive functions.

The third cognitive ability we investigated was working memory. Although we observed a significant correlation between working memory and recognition accuracy, this ability did not directly predict recognition accuracy or adaptation when unaccented SRTs and vocabulary score were also included in our regression analysis. However, working memory did have an indirect relationship with recognition accuracy, mediated by vocabulary knowledge, in our path analysis model. Working memory may therefore support recognition of accented speech via other cognitive abilities (in this case, vocabulary knowledge), as observed in speech reading (Lyxell and Ronnberg, 1989) and perceptual adaptation to distorted visual input (Kennedy *et al.*, 2009). Other studies investigating working memory and perceptual adaptation to unfamiliar speech have produced mixed results: although working memory is the most reliable predictor of recognition of speech in background noise, this is not a wholly consistent finding (Akeroyd, 2008), and indeed we did not observe a correlation between working memory and unaccented SRTs in our study. Janse and Adank (2012) found that working memory predicts overall recognition accuracy of novel-accented speech in older adults (possibly reflecting greater individual variation in an older population), but no other study has observed this, in foreign-accented (Gordon-Salant *et al.*, 2013), frequency compressed (Ellis and Munro, 2013), or noise-vocoded (Erb *et al.*, 2012) speech.

Our findings, together with current evidence, suggest therefore that working memory does not always play a prominent role in perceptual adaptation to, or recognition of, unfamiliar speech. Furthermore, our effects were small even in a



sample of 100 participants. Studies with smaller samples, and particularly in a young, clinically normal population, may therefore be underpowered to detect such small effects. However, another explanation is also possible. The working memory test used in this study (Romberg *et al.*, 1989) relies specifically on lexical recall, and responses are scored as incorrect if participants recall the correct semantic concept, but not the exact lexical item (e.g., “gun” instead of “pistol”). An overlap with the abilities required for the vocabulary knowledge test (that is, robust mapping between lexical items and semantic concepts) could therefore account for the mediation effect observed in our data.

The present study measured two important aspects of perceptual adaptation to accented speech—recognition accuracy and adaptation (that is, overall performance and changes in performance over time). The results from our regression analyses suggest that different cognitive abilities are involved in these different aspects of adaptation (executive function for amount and rate of adaptation; vocabulary knowledge and, to a lesser extent, working memory, for recognition accuracy). Nevertheless, it should be noted that our measures of recognition accuracy and adaptation rate were significantly correlated, and so differences between these two measures should be interpreted with caution. However, no such correlation was observed between recognition accuracy and adaptation amount, and so we can assume that these measures do indeed reflect different abilities.

## V. CONCLUSION

The present study evaluated the contribution of cognitive ability to perceptual adaptation to accented speech. Results suggest a prominent role for inhibition in perceptual adaptation, and for vocabulary knowledge in overall recognition accuracy. Recognition accuracy was indirectly supported by working memory, via vocabulary knowledge, which suggests that working memory may play a less prominent role in successful recognition of accented speech. Our study is the first to relate inhibition to perceptual adaptation to unfamiliar speech, and substantiates existing evidence that top-down processing, particularly executive function, is important for adapting to speech in adverse listening conditions. However, further investigations may help to discern the exact role of executive function and vocabulary knowledge in perceptual adaptation to accented speech.

## ACKNOWLEDGMENTS

This work was funded by a Biotechnology and Biological Sciences Research Council research studentship awarded to B.B. and by The University of Manchester. The authors thank Stuart Rosen for supplying the adaptive staircase program and the speech-shaped noise.

<sup>1</sup>A one-way ANOVA (5 levels: training group) of SRTs for the accented speech revealed no effect of training group:  $F(4, 95) = 0.23$ ,  $p = 0.92$ ,  $\eta_p^2 = 0.01$ .

<sup>2</sup>It should be noted that starting values did not significantly contribute to the variance in adaptation when included in our regression models: adaptation amount:  $\beta = 0.09$ ,  $t = 0.86$ ,  $p = 0.391$ ; adaptation rate:  $\beta = -0.11$ ,  $t = 1.05$ ,  $p = 0.298$ .

- Adank, P., and Janse, E. (2010). “Comprehension of a novel accent by young and older listeners,” *Psychol. Aging* **25**, 736–740.
- Akeroyd, M. A. (2008). “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.* **47**, S53–S71.
- Amitay, S. (2009). “Forward and reverse hierarchies in auditory perceptual learning,” *Learn. Percept.* **1**, 59–68.
- Baker, R., and Rosen, S. (2001). “Evaluation of maximum-likelihood threshold estimation with tone-in-noise masking,” *Br. J. Audiol.* **35**, 43–52.
- Baron, R. M., and Kenny, D. A. (1986). “The moderator mediator variable distinction in social psychological-research—Conceptual, strategic, and statistical considerations,” *J. Pers. Soc. Psychol.* **51**, 1173–1182.
- Boersma, P., and Weenink, D. (2012). “Praat: Doing phonetics by computer (version 5.3.05) [computer program],” [www.praat.org](http://www.praat.org) (Last viewed February 19, 2012).
- Borovsky, A., Elmana, J. L., and Fernald, A. (2012). “Knowing a lot for one’s age: Vocabulary skill and not age is associated with anticipatory incremental sentence interpretation in children and adults,” *J. Exp. Child Psychol.* **112**, 417–436.
- Bradlow, A. R., and Bent, T. (2008). “Perceptual adaptation to non-native speech,” *Cognition* **106**, 707–729.
- Brouwer, S., Mitterer, H., and Huettig, F. (2012). “Can hearing puter activate pupil? Phonological competition and the processing of reduced spoken words in spontaneous conversations,” *Q. J. Exp. Psychol.* **65**, 2193–2220.
- Clarke, C. M., and Garrett, M. F. (2004). “Rapid adaptation to foreign-accented English,” *J. Acoust. Soc. Am.* **116**, 3647–3658.
- Cristia, A., Seidl, A., Vaughn, C., Schmale, R., Bradlow, A. R., and Floccia, C. (2012). “Linguistic processing of accented speech across the lifespan,” *Front. Psychol.* **3**, 479.
- Davis, M. H., Johnsrude, I. S., Hervais-Adelman, A., Taylor, K., and McGettigan, C. (2005). “Lexical information drives perceptual learning of distorted speech: Evidence from the comprehension of noise-vocoded sentences,” *J. Exp. Psychol. Gen.* **134**, 222–241.
- Dupoux, E., and Green, K. (1997). “Perceptual adjustment to highly compressed speech: Effects of talker and rate changes,” *J. Exp. Psychol. Hum. Percept. Perform.* **23**, 914–927.
- Eisner, F., and McQueen, J. M. (2005). “The specificity of perceptual learning in speech processing,” *Percept. Psychophys.* **67**, 224–238.
- Ellis, R. J., and Munro, K. J. (2013). “Does cognitive function predict frequency compressed speech recognition in listeners with normal hearing and normal cognition?,” *Int. J. Audiol.* **52**, 14–22.
- Erb, J., Henry, M. J., Eisner, F., and Obleser, J. (2012). “Auditory skills and brain morphology predict individual differences in adaptation to degraded speech,” *Neuropsychologia* **50**, 2154–2164.
- Erb, J., Henry, M. J., Eisner, F., and Obleser, J. (2013). “The brain dynamics of rapid perceptual adaptation to adverse listening conditions,” *J. Neurosci.* **33**, 10688–10697.
- Faul, F., Erdfelder, E., Buchner, A., and Lang, A.-G. (2009). “Statistical power analyses using G\*Power 3.1: Tests for correlation and regression analyses,” *Behav. Res. Methods* **41**, 1149–1160.
- Golomb, J. D., Peelle, J. E., and Wingfield, A. (2007). “Effects of stimulus variability and adult aging on adaptation to time-compressed speech,” *J. Acoust. Soc. Am.* **121**, 1701–1708.
- Gordon-Salant, S., Yeni-Komshian, G. H., Fitzgibbons, P. J., Cohen, J. I., and Waldroup, C. (2013). “Recognition of accented and unaccented speech in different maskers by younger and older listeners,” *J. Acoust. Soc. Am.* **134**, 618–627.
- Gordon-Salant, S., Yeni-Komshian, G. H., Fitzgibbons, P. J., and Schurman, J. (2010). “Short-term adaptation to accented English by younger and older adults,” *J. Acoust. Soc. Am.* **128**, EL200–EL204.
- Halliday, L. F., Moore, D. R., Taylor, J. L., and Amitay, S. (2011). “Dimension-specific attention directs learning and listening on auditory training tasks,” *Att. Percept. Psychophys.* **73**, 1329–1335.
- Hervais-Adelman, A., Davis, M. H., Johnsrude, I. S., and Carlyon, R. P. (2008). “Perceptual learning of noise vocoded words: Effects of feedback and lexicality,” *J. Exp. Psychol. Hum. Percept. Perform.* **34**, 460–474.

- Huyck, J. J., and Johnsrude, I. S. (2012). "Rapid perceptual learning of noise-vocoded speech requires attention," *J. Acoust. Soc. Am.* **131**, EL236–EL242.
- Institute of Electrical and Electronics Engineers (1969). "IEEE recommended practices for speech quality measurements," *IEEE Trans. Aud. Electroacoust.* **17**, 227–246.
- Janse, E. (2009). "Processing of fast speech by elderly listeners," *J. Acoust. Soc. Am.* **125**, 2361–2373.
- Janse, E. (2012). "A non-auditory measure of interference predicts distraction by competing speech in older adults," *Aging Neuropsychol. Cog. J. Norm. Dysfunct. Dev.* **19**, 741–758.
- Janse, E., and Adank, P. (2012). "Predicting foreign-accent adaptation in older adults," *Q. J. Exp. Psychol.* **65**, 1563–1585.
- Kamphaus, R. W. (2005). *Clinical Assessment of Child and Adolescent Intelligence*, 2nd ed. (Springer Science, New York), pp. 359–417.
- Kennedy, K. M., Rodrigue, K. M., Head, D., Gunning-Dixon, F., and Raz, N. (2009). "Neuroanatomical and cognitive mediators of age-related differences in perceptual priming and learning," *Neuropsychology* **23**, 475–491.
- Koelewijn, T., Zekveld, A. A., Festen, J. M., Ronnberg, J., and Kramer, S. E. (2012). "Processing load induced by informational masking is related to linguistic abilities," *Int. J. Otolaryngol.* **2012**, 865731.
- Lambert, W. E., Hodgson, R. C., Gardner, R. C., and Fillenbaum, S. (1960). "Evaluational reactions to spoken languages," *J. Abnorm. Soc. Psychol.* **60**, 44–51.
- Lyxell, B., and Ronnberg, J. (1989). "Information-processing skill and speech-reading," *Br. J. Audiol.* **23**, 339–348.
- Mattys, S. L., Davis, M. H., Bradlow, A. R., and Scott, S. K. (2012). "Speech recognition in adverse conditions: A review," *Language Cognit. Processes* **27**, 953–978.
- Maye, J., Aslin, R. N., and Tanenhaus, M. K. (2008). "The weckud wetch of the wast: Lexical adaptation to a novel accent," *Cognit. Sci.* **32**, 543–562.
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., and Wager, T. D. (2000). "The unity and diversity of executive functions and their contributions to complex 'frontal lobe' tasks: A latent variable analysis," *Cognit. Psychol.* **41**, 49–100.
- Norris, D., McQueen, J. M., and Cutler, A. (2003). "Perceptual learning in speech," *Cognit. Psychol.* **47**, 204–238.
- Pichora-Fuller, M. K., and Singh, G. (2006). "Effects of age on auditory and cognitive processing: Implications for hearing aid fitting and audiologic rehabilitation," *Trends Amplif.* **10**, 29–59.
- Plomp, R., and Mimpen, A. M. (1979). "Speech-reception threshold for sentences as a function of age and noise-level," *J. Acoust. Soc. Am.* **66**, 1333–1342.
- Preacher, K. J., and Hayes, A. F. (2004). "SPSS and SAS procedures for estimating indirect effects in simple mediation models," *Behav. Res. Methods Instrum. Comput.* **36**, 717–731.
- Ronnberg, J. (2003). "Cognition in the hearing impaired and deaf as a bridge between signal and dialogue: A framework and a model," *Int. J. Audiol.* **42**, S68–S76.
- Ronnberg, J., Lyxell, B., Arlinger, S., and Kinnefors, C. (1989). "Visual evoked-potentials: Relation to adult speechreading and cognitive function," *J. Speech Hear. Res.* **32**, 725–735.
- Ronnberg, J., Rudner, M., Foo, C., and Lunner, T. (2008). "Cognition counts: A working memory system for ease of language understanding (ELU)," *Int. J. Audiol.* **47**, S99–S105.
- Samuel, A. G., and Kraljic, T. (2009). "Perceptual learning for speech," *Att. Percept. Psychophys.* **71**, 1207–1218.
- Schaie, K. W., Willis, S. L., and O'Hanlon, A. M. (1994). "Perceived intellectual performance change over seven years," *J. Gerontol.* **49**, P108–P118.
- Shrout, P. E., and Bolger, N. (2002). "Mediation in experimental and nonexperimental studies: New procedures and recommendations," *Psychol. Methods* **7**, 422–445.
- Singer, T., Verhaeghen, P., Ghisletta, P., Lindenberger, U., and Baltes, P. B. (2003). "The fate of cognition in very old age: Six-year longitudinal findings in the Berlin Aging Study (BASE)," *Psychol. Aging* **18**, 318–331.
- Sommers, M. S., and Danielson, S. M. (1999). "Inhibitory processes and spoken word recognition in young and older adults: The interaction of lexical competition and semantic context," *Psychol. Aging* **14**, 458–472.
- Stacey, P. C., and Summerfield, A. Q. (2007). "Effectiveness of computer-based auditory training in improving the perception of noise-vocoded speech," *J. Acoust. Soc. Am.* **121**, 2923–2935.
- Stroop, J. R. (1935). "Studies of interference in serial verbal reactions," *J. Exp. Psychol.* **18**, 643–662.
- Tuinman, A., Mitterer, H., and Cutler, A. (2012). "Resolving ambiguity in familiar and unfamiliar casual speech," *J. Mem. Lang.* **66**, 530–544.
- Wechsler, D. (1958). *The Measurement and Appraisal of Adult Intelligence*, 4th ed. (The Williams and Wilkins Company, Baltimore, MD), pp. 61–117.
- Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence* (Pearson, San Antonio, TX), pp. 54–86.
- Wild, C. J., Yusuf, A., Wilson, D. E., Peelle, J. E., Davis, M. H., and Johnsrude, I. S. (2012). "Effortful listening: The processing of degraded speech depends critically on attention," *J. Neurosci.* **32**, 14010–14021.