



OPEN ACCESS

EDITED BY

Piers Dawes,
The University of Manchester,
United Kingdom

REVIEWED BY

Rebecca E. Millman,
The University of Manchester,
United Kingdom
Kenneth Vaden,
The Medical University of South
Carolina, United States
Kazuo Ueda,
Kyushu University, Japan

*CORRESPONDENCE

Andrew M. Burluson
andrewburluson@u.northwestern.edu

SPECIALTY SECTION

This article was submitted to
Auditory Cognitive Neuroscience,
a section of the journal
Frontiers in Psychology

RECEIVED 01 October 2022

ACCEPTED 23 November 2022

PUBLISHED 08 December 2022

CITATION

Burluson AM and Souza PE (2022)
Cognitive and linguistic abilities
and perceptual restoration of missing
speech: Evidence from online
assessment.
Front. Psychol. 13:1059192.
doi: 10.3389/fpsyg.2022.1059192

COPYRIGHT

© 2022 Burluson and Souza. This is an
open-access article distributed under
the terms of the [Creative Commons
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,
distribution or reproduction in other
forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution
or reproduction is permitted which
does not comply with these terms.

Cognitive and linguistic abilities and perceptual restoration of missing speech: Evidence from online assessment

Andrew M. Burluson* and Pamela E. Souza

Hearing Aid Laboratory, Department of Communication Sciences and Disorders, Northwestern University, Evanston, IL, United States

When speech is clear, speech understanding is a relatively simple and automatic process. However, when the acoustic signal is degraded, top-down cognitive and linguistic abilities, such as working memory capacity, lexical knowledge (i.e., vocabulary), inhibitory control, and processing speed can often support speech understanding. This study examined whether listeners aged 22–63 (mean age 42 years) with better cognitive and linguistic abilities would be better able to perceptually restore missing speech information than those with poorer scores. Additionally, the role of context and everyday speech was investigated using high-context, low-context, and realistic speech corpi to explore these effects. Sixty-three adult participants with self-reported normal hearing completed a short cognitive and linguistic battery before listening to sentences interrupted by silent gaps or noise bursts. Results indicated that working memory was the most reliable predictor of perceptual restoration ability, followed by lexical knowledge, and inhibitory control and processing speed. Generally, silent gap conditions were related to and predicted by a broader range of cognitive abilities, whereas noise burst conditions were related to working memory capacity and inhibitory control. These findings suggest that higher-order cognitive and linguistic abilities facilitate the top-down restoration of missing speech information and contribute to individual variability in perceptual restoration.

KEYWORDS

perceptual restoration, interrupted speech, cognition, linguistic, online assessment

Introduction

When conditions are optimal, understanding speech for normal-hearing listeners is a relatively simple and automatic process. High-fidelity speech information is rapidly transmitted through the peripheral and central auditory systems to the primary auditory cortex, acoustic cues are matched to stored lexical representations, and meaning can

be extracted with little to no conscious reappraisal (Marslen-Wilson and Welsh, 1978). However, in everyday communication, background noise or interruptions are common. Background noise can interfere with the perception of target speech (Gordon-Salant and Fitzgibbons, 2004; Neuman et al., 2010; Smith et al., 2019), and if the interruptions become more intense than the speech itself, segments may be masked entirely. Despite this obfuscation of the speech signal by background noise, many listeners with normal hearing remain able to understand speech relatively well. In difficult conditions, listeners are thought to piece together the remaining speech fragments to “fill in the gaps,” integrating and organizing them perceptually across time. For example, listeners may rely on spectrotemporal cues, such as the fundamental frequency, that are less affected by background noise and can bridge the gaps to assist with perceptual grouping (Li and Loizou, 2007; Oxenham, 2008). This idea is often referred to as “glimpsing,” “dip listening” (Cooke, 2006; Akeroyd, 2008), or “perceptual restoration” when segments of speech information are intentionally absent or removed (Warren, 1970).

This process can be investigated using an interrupted speech paradigm wherein segments of a speech signal are periodically removed, as pioneered by Miller and Licklider (1950). Periodic removal of speech allows for an investigation into which factors may aid in the recovery of the remaining proportion of speech information. Speech intelligibility improves when the periodic interruption is a rectangular burst of broadband noise instead of a silent gap (Powers and Wilcox, 1977; Bashford and Warren, 1987; Bashford et al., 1992; Bologna et al., 2019). This effect is particularly salient when the intensity of the noise burst is greater than the speech signal. The negative signal-to-noise ratio is thought to give the listener the impression of perceptually continuous speech occurring behind the noise, aiding perceptual organization and grouping (Bashford and Warren, 1987; Bashford et al., 1996).

Most perceptual restoration research has emphasized signal-level factors, such as speech spectrotemporal fidelity, rate, and length; interruption length, density, and type; and interruption/signal intensity that affect how much speech information can be restored (Miller and Licklider, 1950; Warren, 1970; Bashford et al., 1996; Başkent et al., 2009; Chatterjee et al., 2010; Jin and Nelson, 2010; Benard et al., 2014; Clarke et al., 2016; Shafiro et al., 2016). Furthermore, most of the perceptual restoration literature focuses on group differences (e.g., age, hearing status) (Başkent, 2010; Kidd and Humes, 2012; Fogerty et al., 2015; Başkent et al., 2016; Bologna et al., 2018; Jaekel et al., 2018). While signal-level factors clearly play a role in the ability to restore missing speech, individuals within the same group still vary substantially in their perceptual restoration ability.

We suggest that individual variability during the perceptual restoration of missing speech information may be driven by

individual differences in higher-order processing abilities, such as cognitive and linguistic abilities. Cognitive abilities differ substantially from one person to the next, and while abilities such as working memory and inhibitory control do tend to vary together within one person, they are not always aligned (e.g., an individual can have high working memory with low inhibitory control) (Carroll and Maxwell, 1979; Boogert et al., 2018). Linguistic abilities also vary across individuals, with lexical knowledge, or vocabulary, increasing with advancing age (Verhaeghen, 2003). The Ease of Language Understanding (ELU) model takes these higher-order processes into account, proposing a model where cognitive and linguistic abilities interact to support degraded speech understanding. The ELU model accounts for cognitive abilities, such as working memory, which allows speech to be temporarily held in an episodic buffer for later reprocessing; and linguistic knowledge, which allows context and vocabulary to identify possible lexical candidates for speech which was not automatically recognized (Rönnberg et al., 2013). This model provides an explanation whereby individuals may restore missing or interrupted speech differently based on their cognitive and linguistic abilities, as follows.

First, the reconstruction of missing speech requires the reprocessing of available speech fragments which are temporarily held in an episodic buffer. Temporarily holding speech fragments tasks a listener's working memory capacity. Some evidence suggests that an individual's working memory capacity mediates the ability to restore missing speech (Benard and Başkent, 2014; Millman and Mattys, 2017; Nagaraj and Magimairaj, 2017) while other data are less definitive (Nagaraj and Knapp, 2015; Shafiro et al., 2015; Bologna et al., 2018). Second, reprocessing is informed by a listener's lexical knowledge and how quickly that information can be accessed to accurately identify lexical candidates when filling the gap. Current literature indicates that lexical knowledge (i.e., vocabulary) plays a role during perceptual restoration (Benard and Başkent, 2014; Nagaraj and Magimairaj, 2017). However, existing literature has not captured an aspect of speech perception that may be important during perceptual restoration: lexical access speed, or the rate at which stored lexical representations can be activated or matched by speech information. Third, irrelevant information, such as unlikely lexical candidates, noise-burst interruptions, and other cognitive processes that may be competing for attention must be inhibited, which relies on a listener's inhibitory control. Current evidence suggests that an individual's inhibitory control is predictive of his/her ability to perform other degraded speech recognition tasks, such as speech-in-noise (Dey and Sommers, 2015; Dryden et al., 2017; Stenbäck et al., 2021; Perron et al., 2022). Bologna et al. (2018) investigated inhibitory control using interrupted speech and found null results for both younger and older adults, albeit at a very difficult signal-to-noise ratio. Last, working memory reprocessing, lexical

processing, and inhibitory control require processing time to complete. Thus, these abilities depend on processing speed, or the rate at which cognitive tasks are completed by an individual (Salthouse, 1992; Morrison and Gibbons, 2006; Rozas et al., 2008). Processing speed has shown predictive value in previous research on degraded speech recognition (Ellis et al., 2016; Dryden et al., 2017; Yumba, 2017) and perceptual restoration (Bologna et al., 2018), but data exist only for young normal-hearing and older hearing-impaired adults.

Taken together, we predict that listeners who have a higher working memory capacity, greater lexical knowledge, faster lexical access speed, better inhibitory control, and faster processing speed will be more successful when restoring missing speech information, especially for high context, predictable sentences. Building from previous work, we used a periodically interrupted speech paradigm to force listeners into an explicit processing loop as outlined in the ELU model (Rönnberg et al., 2013), and we separately measured the cognitive and linguistic processes supporting explicit processing. To further explore the role of lexical processing during perceptual restoration, we chose to include both high- and low-context sentences in addition to a sentence set that resembles everyday speech. Because in-person testing capacity was restricted as a result of the COVID-19 pandemic, this experiment was conducted using online assessments for listeners ranging from young to middle-aged adults.

Materials and methods

Participants

Prior to data collection an *a priori* power analysis was performed based on the relationship between perceptual restoration and cognitive data from Nagaraj and Magimairaj (2017). For a medium effect size of 0.3, an alpha level of 0.05, a power level of 0.9, and four predictors, the projected sample size necessary was 57 participants. Sixty-three participants (22 males, 36 females, five other or prefer not to answer) completed this experiment (Age range = 22–63 years, Mean = 42.0 years, SD = 12 years); they represented an age range captured in only a small set of perceptual restoration data (Millman and Mattys, 2017). To be eligible, participants needed to self-report normal hearing and cognitive status, speak English as their primary language, be between 18 and 65 years of age, and be a current resident of the United States. Because participation was virtual, hearing thresholds were not assessed. The Institutional Review Board of Northwestern University approved the study, all participants signed an informed consent form on the secure data collection platform REDCap (Harris et al., 2009), and participants were compensated at an hourly rate for taking part in the study.

Stimuli

Speech stimuli consisted of three sentence sets: the Revised Speech in Noise (RSPIN) low- and high-context sentences (Bilger et al., 1984) and the Perceptually Robust English Sentence Test: Open Set (PRESTO) (Tamati et al., 2013). The RSPIN sentences were designed to determine the role of top-down and bottom-up processes during speech recognition and were selected from a corpus of 200 sentences that were highly predictable (e.g., “the witness took a solemn *oath*”), where top-down resources can inform final word choice, or 200 sentences that were unpredictable but syntactically correct (e.g., “he has a problem with the *oath*”) which relies more on the fidelity of the bottom-up signal to the auditory cortex. Following Jenstad and Souza (2007), the entire sentence was scored (see Section “General procedure”). RSPIN sentences were produced by a male talker. High context sentences had an average of 5.1 content words per sentence, and low context sentences had an average of 4.8 content words per sentence. The PRESTO is a high-variability sentence set designed to be sensitive to individual differences and is thought to access both central cognitive and perceptual abilities during speech recognition, including current theories of lexical organization and automatic encoding of lexical components. The PRESTO sentence set is balanced for talker gender, number of keywords (average of 4.2 content words per sentence), word frequency, and word familiarity.

Both silent gap sentences and noise burst sentences were constructed using a common method for interrupted speech stimuli development which is known to induce perceptual restoration. This method also avoids both floor and ceiling effects, as follows.

All sentences

Six sentence conditions consisting of sixty sentences each were tested: two interruption conditions (silent gap versus noise burst) by three sentence conditions (RSPIN low context, RSPIN high context, and PRESTO), resulting in 240 RSPIN sentences and 120 PRESTO sentences. First, the 360 sentences were gated with a 50% duty cycle square wave at a rate of 2.0 Hz using a custom MATLAB R2020a script, creating interrupted speech stimuli with alternating 250 ms segments of speech and silence. Second, a separate set of 360 noise-burst stimuli were created where the noise bursts aligned with the silent segments of the interrupted speech stimuli. The speech-shaped noise bursts were generated using the combined Fourier transform of all 360 sentences, where the phases of all spectral components were randomized before being converted back into the time domain using an inverse Fourier transform. The overall lengths of the noise-burst stimuli were the same as the overall lengths of the interrupted speech segments because the noise bursts would later be interleaved with the

interrupted speech segments (i.e., creating alternating 250 ms segments of speech and noise bursts). To minimize spectral splatter and distortion, 10 ms cosine on- and off-ramps were applied to both the remaining interrupted speech segments and noise burst stimuli. The RMS of the interrupted speech stimuli and the noise burst stimuli were normalized. Because the amount of speech information restored improves with the addition of a noise burst when the noise burst is louder than the remaining speech segments (Bashford et al., 1996), the level of the noise burst stimuli was raised by 10 dB (–10 dB SNR) relative to all interrupted speech segments. The RMS of interrupted speech segments and the noise bursts was then normalized. By processing the stimuli this way, the level of the speech is always the same for both the silent-gap and noise-burst sentences, while the level of the noise burst will be 10 dB higher than the speech for the noise-burst sentences after processing.

Silent gap sentences

Silent gap sentences consisted of half of the original 360 sentences (120 RSPIN high- and low-context and 60 PRESTO sentences). The preceding procedure resulted in a set of silent gap interrupted speech stimuli with alternating segments of 250 ms of speech and 250 ms of silence with 10 ms cosine on- and off-ramps with a normalized RMS; no further signal processing was required.

Noise burst sentences

For the remaining half of the sentences (120 RSPIN high- and low-context and 60 PRESTO sentences), a periodic noise burst filled the silent gap. To do this, the interrupted speech segment stimuli and the noise-burst stimuli were added linearly to one another, including their individual 10 ms on- and off-ramps eliminating distortion and spectral splatter.

General procedure

Testing was carried out using the online recruitment and experimental testing platforms Prolific and Gorilla, respectively. Pre-screening criteria (see Section “Participants”) was entered into Prolific to identify potential eligible participants, who, after indicating interest, were directed to REDCap (Harris et al., 2009), a secure data collection platform, to complete the consent form, enter demographic data, complete a brief hearing health questionnaire, and to complete the Speech and Spatial Qualities questionnaire (see Section “Questionnaires”). From there, participants were directed to the experimental platform, Gorilla, where they completed cognitive and linguistic tasks (see Sections “Cognitive tasks and Linguistic task”), a headphone screening task, and the interrupted speech task (see Section “Stimuli”).

Questionnaires

First, participants completed a simple demographic questionnaire, followed by a hearing health questionnaire that included self-report of hearing loss and cognitive or memory concerns (participants were excluded if they answered “yes”). Last, participants completed the 49-item Speech and Spatial Qualities of Hearing questionnaire (SSQ) in which participants self-assessed their hearing ability in specific contexts and situations on a numerical scale of 0–10 (Gatehouse and Noble, 2004). Questions address self-perceived function in three domains: speech hearing (“SSQ—Speech”), spatial hearing (“SSQ—Spatial”), and quality of hearing (“SSQ—Quality”). Participants were asked to rate their ability to hear and understand speech in different settings (speech hearing domain), their ability to listen in different environments, which includes distance, direction, and movement (spatial hearing domain), and their perceived abilities for everyday sounds, including music listening, ease of listening, clarity, and naturalness of sound (quality of hearing domain).

Cognitive tasks

To assess listeners’ working memory capacity, inhibitory control, and processing speed, participants completed several automated, virtual assessments in the visual modality: the Reading Span Task (RST; complex working memory capacity), the Digit Span Forward and Backward (DST; simple working memory capacity), the Stroop Task (Stroop; processing speed and inhibitory control), and the Flanker Task (Flanker; processing speed and inhibitory control).

Reading span task

The reading span task (RST) is a task that measures a listener’s complex working memory capacity, or the simultaneous storage and reprocessing of complex information, requiring additional processing beyond simple repetition or reversal of information (see Section “Digit span forward and backward”). The current version of the RST was described by Rönneberg et al. (1989), which was modified from the original version first introduced by Daneman and Carpenter (1980). The current version was modified so that the assessment could be completed virtually and without supervision. In this task, listeners were asked to first read and comprehend sentences presented on a screen and to determine whether or not the sentence makes sense. Half of the sentences were absurd (e.g., “The fish drove a car”) and the other half were normal sentences (e.g., “The ball bounced away”). Each content word and any accompanying articles (e.g., “the ball” or “a car”) were presented sequentially on the screen each for 800 ms. Listeners were then asked to respond “yes” by pressing a button on the screen if the sentence made sense or “no” if the sentence was absurd.

If listeners did not respond within 3,000 ms, the program advanced automatically. Participants were presented with 2–5 sentences per sequence. Listeners were then asked to recall either the first content word or the last content word from each sequence. They were not made aware beforehand whether they will be expected to recall the first or the last word, and thus must maintain both streams of information simultaneously. Using their keyboard, listeners typed their content word responses into a box on the screen and the number of correctly recalled words (not in correct serial order) out of the number of possible words was scored. Because participants were not supervised during this task, practice trials with feedback were provided. First, participants practiced only responding whether or not the sentence made sense. Next, they practiced recalling the first words of a two-sentence sequence, then the last words of a two-sentence sequence. Last, they practiced responding by recalling either the first or the last words of a two-sentence sequence before beginning the actual task. The percent correct of first or last words correctly recalled in any order (“*RST Percent Correct*”) reflects a participant’s complex working memory capacity.

Digit span forward and backward

The digit span task represents a traditional neuropsychological measure of a listener’s short-term memory (digit forward), such as the storage of a phone number (Jones and Macken, 2015), and simple working memory capacity (digit backward). Digit span backwards requires that the participant store and later invert the serial presentation of numerical information, similar to the storage and reprocessing of information during more demanding working memory tasks like the RST. Digit span forward and backward then may represent reduced processing demands compared to the RST (Daneman and Merikle, 1996) or different processes of working memory, with digit span forward and backward tapping into the simple rehearsal of visual stimuli during working memory and RST tapping into more complex rehearsal and reprocessing of visual information in the current study (Millman and Mattys, 2017). However, these complex working memory tasks correlate weakly with digit span backwards and the role of the digit span task as an assessment of working memory has been questioned (Hilbert et al., 2015). The digit span forward and backward task was chosen in addition to the RST to assess a range of memory capacities, from simple to complex, and their relationship to restoration of missing speech across participants. The current digit memory test was designed and revised by Turner and Ridsdale (2004). Participants were presented with a sequence of 2–9 digits and were afterwards asked to type them into the computer, either in the same order for digit span forward, or in reverse order for digit span backward. Each digit was presented on the screen for 1,000 ms. If participants typed in an incorrect response for both trials of a given sequence length, the task would end. Prior to administration of digit span forward and

digit span backward, participants had two practice trials in which they received feedback for each task. Percentiles were calculated from norms and were based on the total number of correct trials for digit span forward and digit span backward together (“*DST Percentile*”) (Turner and Ridsdale, 2004).

Stroop task

The Stroop task measures a participant’s inhibitory control, or their ability to suppress task-irrelevant information. The ability to inhibit irrelevant verbal information, such as unlikely lexical candidates, may allow some listeners to restore more missing speech than others. In the Stroop task, participants named color words (*W*, 25 items) (e.g., “blue”), color hues of “XXXX” to eliminate any reading component (*C*, 25 items), and color words printed in an incongruent color hue (*CW*, 25 items) (e.g., “blue” written in green ink). For the incongruent trials, the participant was asked to name the color of the ink that the word is printed in, not the word itself. The task-naming color words-captures processing speed in milliseconds [*Stroop Processing Speed (ms)*], while the final task captures a participant’s interference score, with higher interference scores indicating reduced inhibitory control and poorer performance (Jensen, 1965). This assessment was based on the method developed by Golden (1976); however, rather than the number of items completed within a specified time limit, each participant completed the same number of items and correct/incorrect and reaction time for each item were captured. For each item, the participant pressed a key on their keyboard that corresponded with the first letter of the color (e.g., “b” for blue). Reminders for the keys were present on the screen. Interference was calculated as the ratio of the average time in milliseconds to correctly identify a *CW* trial divided by the average time taken to correctly identify a *C* trial (i.e., CW/C), a method common in neuropsychology literature [*Stroop Interference (ms)*] (Lansbergen et al., 2007; Scarpina and Tagini, 2017).

Flanker task

The Flanker task measures a participant’s response inhibition, or the ability to suppress responses that are irrelevant or inappropriate for a given task. The Flanker task requires participants to inhibit irrelevant non-verbal information, such as noise bursts, which may allow some listeners perform better on some perceptual restoration tasks than others. During this task, participants completed a computerized version of the Eriksen flanker task (Eriksen and Eriksen, 1974). During this task, participants were presented with five black arrows against a white background and were asked to press a key (“e” for left-facing arrows and “i” for right-facing arrows) to indicate the direction of the arrow in the middle. Participants were asked to respond as quickly and as accurately as possible. There was no time limit for responding on each trial. Half of the 90 items were congruent (e.g., >>>> or <<<<<) and half were incongruent (e.g., >><>> or <<><<). The interstimulus

interval was 750 ms. Before the scored trials, participants had eight practice trials in which they received feedback. Reaction time for congruent and incongruent items were captured as well as task accuracy. Interference was calculated by subtracting the mean reaction time for correct congruent items from the mean reaction time for correct incongruent items in milliseconds [“*Flanker Interference (ms)*”] (Sanders et al., 2018).

Linguistic task

To assess listeners’ lexical access accuracy and lexical access speed, participants completed an automated virtual assessment in the visual modality, the Lexical Test for Advanced Learners of English (LexTALE; lexical knowledge and lexical access speed).

Lexical test for advanced learners of English

The English version of the LexTALE task (Lemhöfer and Broersma, 2012) estimates English vocabulary size (i.e., lexical knowledge) and the speed at which lexical decision-making occurs (i.e., lexical access speed). This measure was originally developed to assess lexical knowledge for intermediate to advanced learners of English as a second language. However, participants who speak English as their first language do not necessarily produce ceiling effects (Lorette and Dewaele, 2015) because factors such as age can influence lexical knowledge over time (Keuleers et al., 2015). Participants were presented with 60 items, 40 of which are real English words and 20 of which are orthographically permissible, pronounceable non-words. Participants were asked to press the “j” key if the word is a real word or the “k” key if it was a non-word and to respond as quickly and as accurately as possible. Reminders for the keys were present on the screen. Participants had 2,000 ms to respond before the program automatically advanced, scoring the missed item as incorrect. Participants did not receive practice trials or feedback prior to task administration. Lexical knowledge was the average of correct responses for real English words and non-words [“*LexTALE Non-word Accuracy (ms)*”] while lexical access speed was measured using the reaction time (“*LexTALE Word RT*”) of correctly identified real words.

Screening task

Participants were asked to wear headphones and to set the volume on their computer to a “loud, but not uncomfortable” level while listening to a recorded excerpt from the Discourse Comprehension Test (Brookshire and Nicholas, 1984). Listeners were also asked to complete a headphone screening procedure to ensure headphone use [see Woods et al. (2017) for more detail]. Briefly, the headphone test required the listener to listen to three tones and pick the softest one out of three correctly at least 4/6 times. Over a loudspeaker setup (e.g., laptop),

one of the three tone presentations suffers from destructive interference resulting from two tones presented out of phase at each loudspeaker, making it difficult to differentiate from the tone that is 6 dB below the standard tone. With headphones, the phase differences do not result in destructive interference, making one of the three tones easier to pick out as the softest. Failing the headphone screening twice resulted in exclusion.

Interrupted speech task

Participants listened to and practiced typing in uninterrupted sentences, followed by those same sentences interrupted by both silent gaps and noise bursts. Feedback was not provided. Participants then listened to the experimental interrupted stimuli. The order of the silent gap sentences and the noise burst sentences were blocked and counterbalanced to prevent order effects. Within each (silent gap or noise burst) block, sentences were not blocked by sentence type and were presented in a random order. After one RSPIN or PRESTO sentence was presented, listeners were asked to type in what they heard into a box on the computer screen. The number of keywords correctly identified was scored using Autoscore (Borrie et al., 2019). Autoscore is an open-source tool for scoring listener transcripts, where the researcher specifies the scoring rule and under which circumstances that rule should be applied. Strict criterion were applied in this experiment. Only the double-letter rule was applied, which scores a word as correct if a double letter is omitted within a word (e.g., “atack” is considered correct for “attack”). Additionally, a custom acceptable spelling list was created that included common misspellings of all keywords in the RSPIN and PRESTO sentences including the following: single letter transpositions within a single word during typing, inclusion/omission of an apostrophe for keywords with a contraction, and any entry of a double space (e.g., spacebar was accidentally hit twice). Traditionally, only the last word of the RSPIN is scored; however, we were interested in how participants restored speech across the entire interrupted sentence, not just the word in the final position. Therefore, content words across the entire sentence were scored using the same method and number of keywords as Jenstad and Souza (2007).

Statistical approach

All data were analyzed using the open source RStudio statistical program version 4.0.5 (R Core Team, 2013), using the *tidyverse* library (Wickham et al., 2019) including the library *dplyr* for data manipulation (Wickham et al., 2022) prior to statistical analysis. The library *ggplot2* was also used for data visualization and figure preparation (Wickham, 2016). For the analysis of variance, the library *rstatix* was utilized

(Kassambara, 2021). For the linear models, the libraries *MASS* and *lmtest* were used to assess homoscedasticity and the distribution of residuals (Venables and Ripley, 2002; Zeileis and Hothorn, 2002). First, outliers in the data were identified and adjusted, followed by descriptive analysis for participant data, cognitive and linguistic measures, and interrupted speech conditions. Next, an analysis of variance (ANOVA) tested for significant differences between the six sentence conditions and Pearson correlations between cognitive and linguistic variables and interrupted sentence conditions were determined. Last, a set of linear regression analyses was performed using cognitive and linguistic variables as predictors for the six sentence conditions.

Results

Prior to analysis, outliers were identified and adjusted, and a fence was determined. All values within any single measure that were outside three times the interquartile range (IQR) were identified as outliers and were adjusted to the nearest fence boundary (i.e., the first or third quartile) to minimize regression toward the mean. Three times the IQR was chosen as a conservative fence in order to avoid unnecessary adjustment given the unsupervised, online nature of the data collected. In total, nine of 1,071 observations across the seventeen reported measures fell outside of the IQR fence and were adjusted. Of the nine adjusted observations, six occurred in the linear models that follow. Descriptive statistics for the 63 participants in this study are presented in Table 1 and results for cognitive and linguistic measures and interrupted speech conditions are available in Table 2. After addressing outliers, measures were normally distributed with skewness and kurtosis under accepted values (Kline, 2015). Participants in this sample performed slightly better but within one standard deviation on the RST compared to existing data (Friedman and Miyake, 2005; Füllgrabe et al., 2015), performed above average on the digit span task with an average percentile score of 72.6 (Turner and Ridsdale, 2004), were consistent with existing Stroop data with regard to reaction time but slightly better with regard to interference scores (Langenecker et al., 2004; Van der Elst et al., 2006), Flanker interference scores were within one standard deviation of existing data (Paap et al., 2020), and participants were highly consistent with published data for English monolinguals on the LexTALE task (Dijkgraaf et al., 2016).

Perceptual restoration differences across experimental conditions

Number of keywords correctly identified across the six sentence conditions were analyzed using an analysis of variance (ANOVA) with the sentence conditions as factor levels and

TABLE 1 Descriptive statistics for participants.

Measures	Mean (SD)	Range	Skew	Kurtosis
Age	42.0 (12.0)	[22, 63]	-0.04	-1.2
Education	15.4 (2.8)	[8, 24]	0.27	0.79
SSQ-speech	8.4 (1.4)	[10, 3]	-1.25	2.11
SSQ-spatial	7.7 (1.5)	[10, 4]	-0.4	-0.48
SSQ-qualities	8.6 (1.2)	[10, 5]	-1	0.39

the percent of keywords correctly identified as the dependent variable. The normality assumption was checked and met using quantile-quantile (Q-Q) plots rather than a Shapiro-Wilk test, as the sample size is greater than 50 participants (D'Agostino, 1971). Levene's test for the homogeneity of variances assumption necessary for the ANOVA was significant, indicating the variances for the six sentence conditions were not equal $F_{(5,372)} = 2.45, p = 0.03$ (Levene, 1960). To account for this violation, a Welch one-way test was used which does not require homogeneity of variance (Moder, 2007).

The perceptual restoration of missing speech information differed by sentence condition (Figure 1), $F_{(5,372)} = 99.7, p < 0.001$. *Post-hoc* pairwise *t*-tests with no assumption of equal variances using a Benjamini-Hochberg correction for multiple comparisons revealed that all pairwise differences between the six conditions were statistically significant ($p < 0.05$) and different from one another, except PRESTO Noise and RSPIN Low Silent conditions ($p = 0.58$) and RSPIN High Silent and RSPIN Low Noise conditions ($p = 0.18$) (Benjamini and Hochberg, 1995).

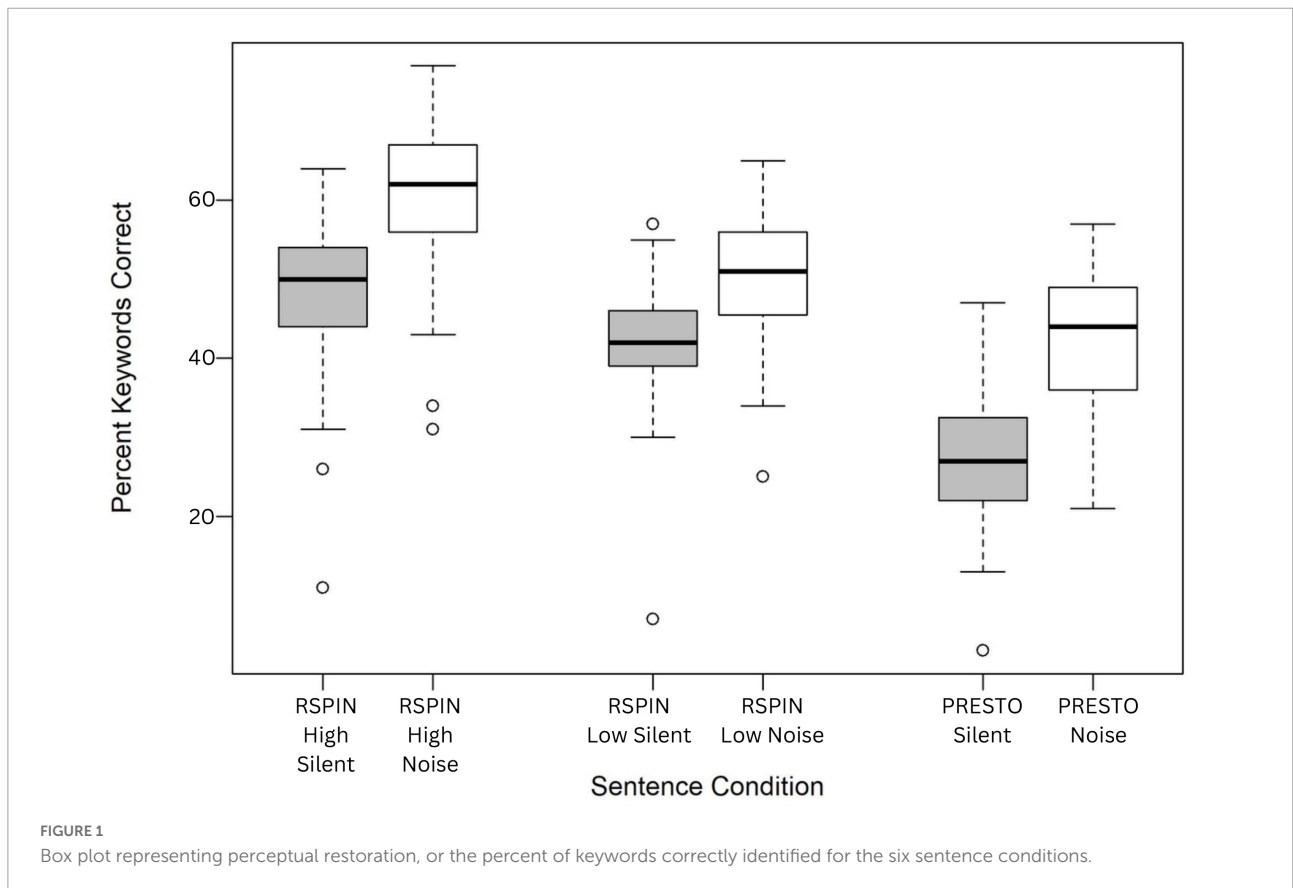
Relationships between perceptual restoration and higher-order, cognitive and linguistic variables

Correlations between the six sentence conditions and cognitive and linguistic variables are presented in Table 3 and Figure 2 (note that *p*-values have not been adjusted for multiple comparisons). Complex working memory capacity measured with the Reading Span Task was moderately correlated with simple working memory measured using the digit span task, a traditional measure working memory thought to be less taxing than complex working memory tasks ($r = 0.28, p = 0.02$). This is consistent with previous research (Daneman and Merikle, 1996) and with similar construct validity between complex working memory, or tasks requiring substantial information storage and reprocessing, and simple working memory, or tasks requiring more straightforward repetition or reversal of information (Lehto, 1996), though the digit span task was not correlated with interrupted speech performance and may not necessarily represent working memory performance (Jones and Macken, 2015). Furthermore, working memory capacity had a moderate,

TABLE 2 Descriptive data for experimental tasks.

	Measures	Mean (SD)	Range	Skew	Kurtosis
Cognitive measures	RST percent correct	73.5 (14.3)	[96.7, 34]	-1.31	1.66
	DST percentile	72.6 (28.6)	[99.9, 0.8]	-1.22	0.42
	Stroop interference (ms)	1.35 (0.26)	[2.29, 0.95]	1.07	1.33
	Stroop processing speed (ms)	804.9 (156.9)	[1,337, 472]	0.84	1.41
	Flanker interference (ms)	41 (23.7)	[100, -12.5]	0.52	0.21
Lexical tasks	LexTALE percent correct	88.7 (9.4)	[100, 66.3]	-1.13	0.45
	LexTALE word RT (ms)	754.3 (125.5)	[1,110, 476]	0.52	0.4
Interrupted speech conditions	RSPIN high silent	48.3 (8.7)	[64, 26]	-0.55	-0.32
	RSPIN high noise	60.1 (9.1)	[77, 31]	-0.97	1.28
	RSPIN low silent	42.4 (5.9)	[57, 30]	0.2	0.01
	RSPIN low noise	50.4 (7.7)	[65, 25]	-0.66	0.84
	PRESTO silent	27.6 (8.3)	[47, 3]	-0.03	0.46
	PRESTO noise	43.1 (8.7)	[57, 21]	-0.3	-0.54

RSPIN, revised speech in noise test; PRESTO, perceptually robust English sentence test open-set; High refers to high context sentences; Low refers to low context sentences; Silent refers to sentences interrupted by a silent gap; and Noise refers to sentences interrupted by a noise burst.



negative correlation with inhibitory control measured using the Flanker task ($r = -0.28, p = 0.02$), which was the only significant correlation with the Flanker task across all measures, making it a weak predictor overall. Lexical processing speed recorded using the LexTALE word reaction time in milliseconds was positively and significantly correlated with inhibitory control measured

using the Stroop Interference score ($r = 0.29, p = 0.02$) and processing speed measured using the Stroop word-only item reaction time, or processing speed ($r = 0.29, p = 0.02$). This result is consistent with both processing speed and the calculation of inhibitory control both relying on reaction time. Working memory measured using the Reading Span Task significantly

TABLE 3 Pearson correlation coefficients across cognitive, linguistic, and interrupted speech measures.

Measures	1	2	3	4	5	6	7	8	9	10	11	12
Cognitive measures												
1 RST percent correct	1											
2 DST percentile	0.28 (0.02)	1										
3 Stroop interference (ms)	0.005 (0.96)	0.003 (0.98)	1									
4 Stroop processing speed (ms)	-0.15 (0.25)	0.07 (0.55)	-0.07 (0.60)	1								
5 Flanker interference (ms)	-0.28 (0.025)	-0.12 (0.38)	0.08 (0.51)	0.13 (0.30)	1							
Linguistic measures												
6 LexTALE percent correct	0.04 (0.74)	-0.10 (0.43)	-0.19 (0.12)	-0.13 (0.29)	-0.15 (0.21)	1						
7 LexTALE word RT (ms)	-0.22 (0.08)	-0.02 (0.86)	0.29 (0.02)	0.29 (0.02)	0.08 (0.51)	-0.18 (0.17)	1					
Interrupted speech measures												
8 RSPIN high silent	0.38 (0.002)	-0.03 (0.82)	-0.12 (0.36)	-0.22 (0.08)	-0.11 (0.40)	0.32 (0.009)	-0.12 (0.37)	1				
9 RSPIN high noise	0.40 (0.001)	-0.05 (0.68)	-0.16 (0.22)	-0.13 (0.32)	-0.01 (0.94)	0.10 (0.39)	-0.18 (0.15)	0.63 (<0.001)	1			
10 RSPIN low silent	0.18 (0.16)	-0.1 (0.43)	-0.27 (0.03)	-0.32 (0.01)	0.001 (0.99)	0.34 (0.006)	-0.13 (0.31)	0.73 (<0.001)	0.47 (<0.001)	1		
11 RSPIN low noise	0.32 (0.01)	-0.04 (0.74)	-0.27 (0.03)	-0.21 (0.11)	-0.04 (0.74)	0.13 (0.30)	-0.22 (0.08)	0.62 (<0.001)	0.83 (<0.001)	0.55 (<0.001)	1	
12 PRESTO silent	0.32 (0.01)	-0.07 (0.59)	-0.06 (0.64)	-0.38 (0.002)	-0.08 (0.56)	0.35 (0.004)	-0.21 (0.09)	0.80 (<0.001)	0.48 (<0.001)	0.80 (<0.001)	0.57 (<0.001)	1
13 PRESTO noise	0.38 (0.002)	0.006 (0.96)	-0.27 (0.03)	-0.26 (0.04)	-0.05 (0.67)	0.26 (0.03)	-0.27 (0.03)	0.78 (<0.001)	0.79 (<0.0001)	0.63 (<0.001)	0.79 (<0.001)	0.67 (<0.001)

Correlation strength is followed by statistical significance, and bolded cells are statistically significant without correction for multiple comparisons ($p < 0.05$).

and positively correlated with five of the six sentence conditions. Lexical knowledge measured using the LexTALE percent correct score correlated with four of the six sentence conditions. Both inhibitory control measured using the Stroop interference score and processing speed measured using the Stroop word-only item reaction time correlated with three of the six sentence conditions. After correcting for multiple comparisons using the Bonferroni method for 24 comparisons (six conditions times four measures of interest), reducing the α level to 0.00208, only RST was significantly correlated with RSPIN High Silent ($r = 0.38, p = 0.002$), RSPIN High Noise ($r = 0.40, p = 0.001$), and PRESTO Noise ($r = 0.38, p = 0.002$).

Some measures had few or no correlations with the interrupted sentence conditions. For example, lexical access speed measured using the LexTALE correctly identified word reaction time in milliseconds correlated only with the PRESTO noise-burst interrupted sentence condition. Simple working memory measured using the digit span task and inhibitory control measured using the Flanker task did not correlate with any of the sentence conditions. These latter three variables were considered poor predictors and were excluded from further analysis. Age was significantly correlated with only the PRESTO silent gap interrupted sentence condition ($r = -0.4, p = 0.001$) and correlated with only the Stroop processing speed cognitive measure ($r = 0.35, p = 0.004$). Age was not significantly correlated with perceptual restoration performance or performance on the cognitive and linguistic measures overall and was excluded from further analysis.

Linear regression analysis was performed using normalized predictors and word recognition percent correct outcome data. Separate models were conducted for each sentence condition. Predictors for the models for the sentence conditions were selected using an *a priori*, hypothesis-driven approach. This approach was informed by the results of Table 3 to minimize Pearson correlation coefficients between predictors during linear model design (Bursac et al., 2008; Hosmer et al., 2013). Last, *a priori* model design was checked against a quantitative approach to minimize the number of predictors while maximizing numerical stability and ease of interpretation [i.e., purposeful selection (Zhang, 2016)]. This approach removes predictors, one by one, from a full, saturated model when their p -values are less than 0.25, unless they are assumed to be related to the hypothesis (Mickey and Greenland, 1989). During this process, predictors such as age, education, the SSQ, DST, and Flanker were not significantly associated with the six interrupted speech conditions and were systematically removed from the model. This method then creates a new, smaller model which can be compared to the saturated model to ensure that the change in coefficients ($\Delta\beta$) is not greater than 20%, which would indicate that these predictors should be added back into the model given their strong adjustment effect. Last, potential interactions among remaining predictors are checked one-by-one and removed if non-significant before goodness of fit (GOF)

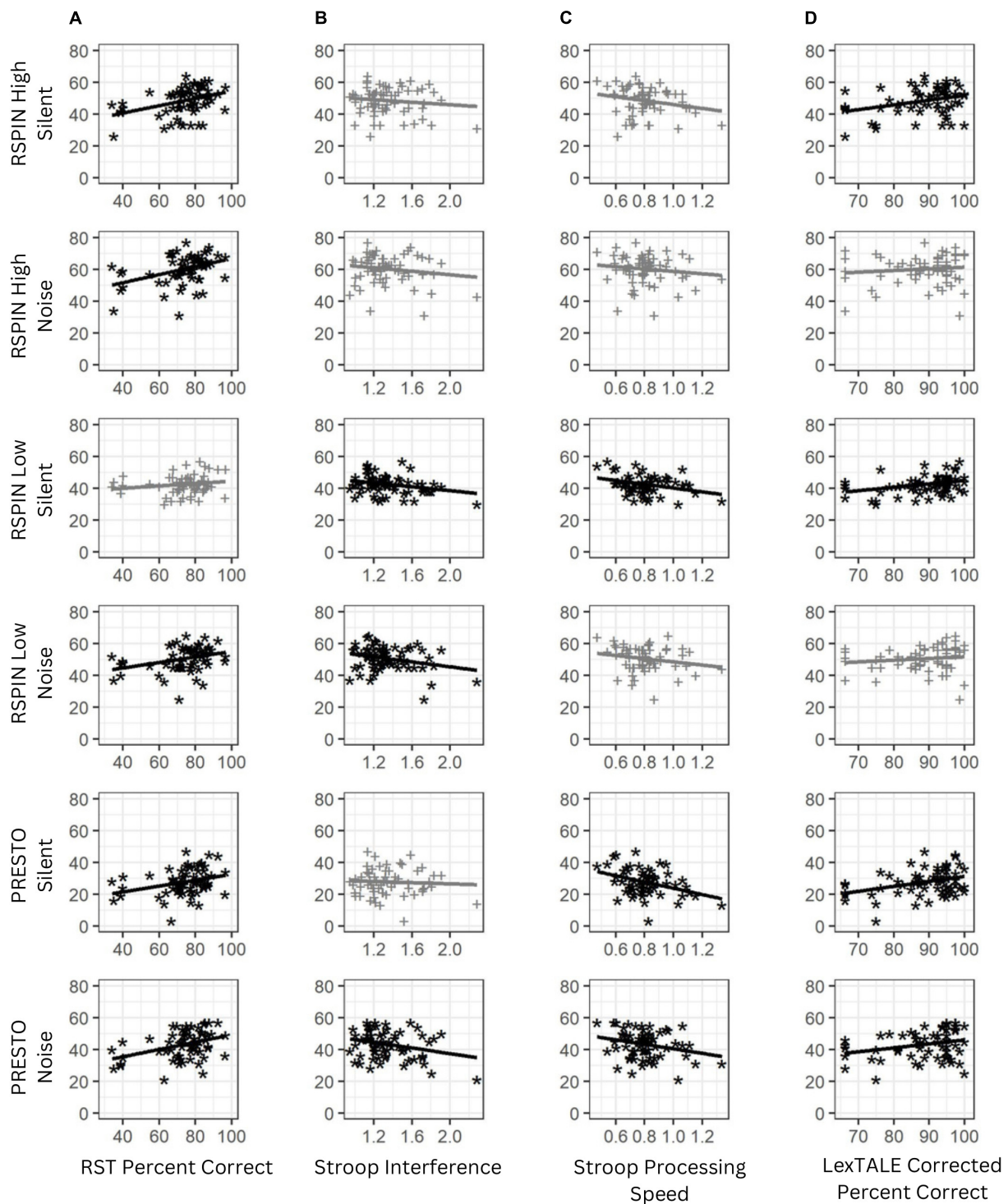


FIGURE 2

Scatter plots of perceptual restoration scores for interrupted speech conditions with working memory capacity measured using the reading span test (RST) (A), inhibitory control measured using the Stroop task (B), processing speed measured using Stroop reaction time (C), and lexical accuracy measured using the LexTALE Percent Correct score (D). Higher scores on RST Percent Correct and LexTALE Percent Correct indicate better performance, while lower scores on Stroop Interference and Stroop Processing Speed indicate better performance. Scatter plots in black with asterisk symbols are statistically significant without correction for multiple comparisons ($p < 0.05$) and scatter plots in gray with plus symbols are not statistically significant. Refer to [Table 3](#) for correlation strength and statistical significance for each of these measures.

is checked visually using plots of residual values versus fitted values and Q-Q plots. This purposeful selection approach was completed for each sentence condition using the percent of words correctly identified as outcome. This process resulted in an overall standard model with the same four predictors used in each model for ease of interpretation: RST Percent Correct, LexTALE Percent Correct, Stroop Interference, and Stroop Processing Speed. Multicollinearity was deemed acceptable among these predictors with the highest variance inflation factor value being 1.06, very close to the minimum value of 1 and well below 2.5, which may indicate multicollinearity, or 10, which is problematic (Mansfield and Helms, 1982; Vittinghoff et al., 2006; Thompson et al., 2017; Johnston et al., 2018). All six models meet the assumption of homoscedasticity necessary for linear model design using a studentized Breusch-Pagan test (Breusch and Pagan, 1979). Three (RSPIN Low Silent, PRESTO Silent, and PRESTO Noise) models met the assumption that the residuals

are normally distributed using a Wilk-Shapiro test of normality, while the remaining three models (RSPIN High Silent, RSPIN High Noise, RSPIN Low Noise) have a non-normal distribution of residuals and should be interpreted with caution.

The overall linear models for all six conditions were statistically significant ($p < 0.05$). All six models are reported without correction for multiple comparisons (see Table 4) as there is not a strong consensus regarding correction when considering multiple separate models. However, a Bonferroni correction for six comparisons reduces the α level to 0.008 and five of the six models remain significant, with only RSPIN High Noise losing significance ($p = 0.01$). Adjusted R^2 values ranged from 13.6 for RSPIN High Noise to 25.5 for PRESTO Silent. Working memory capacity measured using the RST was a significant predictor in five of the six models: RSPIN High Silent ($\beta = 3.01, p = 0.004$), RSPIN High Noise ($\beta = 3.51, p = 0.002$), RSPIN Low Noise ($\beta = 2.28, p = 0.015$), PRESTO Silent ($\beta = 2.16$,

TABLE 4 Linear regression analysis models predicting perceptual restoration by sentence condition.

Condition	Predictors	R ² (%) / Adjusted R ² (%)	F (p-value)	β	Confidence intervals	T-score (p-value)
RSPIN high silent		25.8/20.7	5.05 (0.001)			
	<i>RST percent correct</i>			3.01	1.01–5.01	3.02 (0.004)
	Stroop interference			-0.66	-2.69–1.36	-0.66 (0.51)
	Stroop processing speed			-1.21	-3.23–0.81	-1.2 (0.24)
RSPIN high noise	<i>LexTALE percent correct</i>			2.40	0.36–4.44	2.36 (0.022)
		19.2/13.6	3.45 (0.01)			
	<i>RST percent correct</i>			3.51	1.33–5.69	3.22 (0.002)
	Stroop interference			-1.42	-3.63–0.79	-1.28 (0.2)
RSPIN low silent	Stroop processing speed			-0.69	-2.90–1.51	-0.63 (0.53)
	<i>LexTALE percent correct</i>			0.47	-1.75–2.69	0.42 (0.67)
		26.3/21.3	5.18 (0.001)			
	RST percent correct			0.76	-0.59–2.12	1.13 (0.26)
RSPIN low noise	<i>Stroop interference</i>			-1.45	-2.82–-0.08	-2.12 (0.04)
	<i>Stroop processing speed</i>			-1.67	-3.04–-0.30	-2.44 (0.02)
	<i>LexTALE percent correct</i>			1.47	0.09–2.85	2.14 (0.04)
		21/15.5	3.84 (0.007)			
PRESTO silent	<i>RST percent correct</i>			2.28	0.46–4.1	2.51 (0.015)
	<i>Stroop interference</i>			-2.09	-3.93–-0.25	-2.27 (0.027)
	Stroop processing speed			-1.35	-3.19–0.5	-1.46 (0.15)
	<i>LexTALE percent correct</i>			0.34	-1.52–2.19	0.36 (0.72)
PRESTO noise		30.3/25.5	6.3 (0.0002)			
	<i>RST percent correct</i>			2.16	0.33–4.00	2.36 (0.02)
	Stroop interference			-0.18	-2.04–1.68	-0.19 (0.84)
	<i>Stroop processing speed</i>			-2.48	-4.34–-0.62	-2.67 (0.01)
PRESTO noise	<i>LexTALE percent correct</i>			2.48	0.61–4.35	2.65 (0.01)
		29.8/24.9	6.14 (0.0003)			
	<i>RST percent correct</i>			2.98	1.05–4.92	3.09 (0.003)
	<i>Stroop interference</i>			-2.17	-4.13–-0.21	-2.22 (0.03)
PRESTO noise	Stroop processing speed			-1.74	-3.69–0.22	-1.78 (0.08)
	<i>LexTALE percent correct</i>			1.55	-0.42–3.52	1.58 (0.12)

Significant predictor titles are italicized with bolded statistics ($p < 0.05$).

$p = 0.02$), and PRESTO Noise ($\beta = 2.98, p = 0.003$). Lexical knowledge measured using the LexTALE task was a significant predictor for restoring missing speech for silent conditions, but not noise conditions: RSPIN High Silent ($\beta = 2.40, p = 0.022$), RSPIN Low Silent ($\beta = 1.47, p = 0.04$), and PRESTO Silent ($\beta = 2.48, p = 0.01$). On the other hand, inhibitory control measured using the Stroop task was a significant predictor for the majority of noise burst conditions and one silent gap condition: RSPIN Low Silent ($\beta = -1.45, p = 0.04$), RSPIN Low Noise ($\beta = -2.09, p = 0.027$), and PRESTO Noise ($\beta = -2.17, p = 0.03$). Finally, processing speed significantly predicted perceptual restoration ability in only two of the six models, but in most of the silent gap conditions: RSPIN Low Silent ($\beta = -1.67, p = 0.02$), and PRESTO Silent ($\beta = -2.48, p = 0.01$).

Discussion

The current study was designed to investigate the role of higher-order cognitive and linguistic abilities, such as working memory capacity, lexical knowledge (i.e., vocabulary), lexical access speed, inhibitory control, and processing speed during the perceptual restoration of missing speech information in adults. Of the measures tested, working memory capacity was the most predictive cognitive ability and lexical knowledge was the most predictive linguistic ability during the restoration of missing speech information. The strength of contribution depended on the type of interruption and on sentence material. In the silent gap conditions, a larger set of cognitive and linguistic abilities predicted the restoration of missing speech information. In the noise burst conditions, only working memory capacity and inhibitory control predicted perceptual restoration ability. For high context sentences, working memory capacity and linguistic knowledge predicted the restoration of missing speech, whereas, in the low context and everyday speech conditions, a larger set of cognitive and linguistic abilities predicted perceptual restoration.

Perceptual restoration of missing speech by interruption type and sentence type

In line with previous research, listeners restored more missing speech when sentences were interrupted with noise bursts rather than silent gaps, regardless of sentence type (Miller and Licklider, 1950; Warren, 1970; Powers and Wilcox, 1977; Bashford et al., 1992; Bologna et al., 2019). This result is thought to occur because of gestalt properties of perceptual organization supporting the percept of continuous speech occurring behind the noise bursts, as long as the noise burst itself would be considered an effective masker of the target signal (Warren and Obusek, 1971; Bashford et al., 1992).

The noise bursts may also mask the accidental perception of word boundary that might occur during silent gap interrupted sentences. This occurs when a word is interrupted by a silent gap and that same silent gap is misinterpreted as the end of a word, resulting in the percept of a non-word. However, the noise burst may override this challenging effect by creating illusory continuity and improving degraded speech recognition (Clarke et al., 2016).

Listeners were also able to benefit from sentence context, attaining better scores for RSPIN High sentences when compared to RSPIN Low sentences, a result that is in line with existing perceptual restoration literature (Bashford and Warren, 1987; Bashford et al., 1992; Kidd and Humes, 2012). However, the benefit of context may be limited by the constraints of the RSPIN sentences themselves. Originally, the RSPIN sentences were designed so that only the last word of each sentence would be scored as either correct or incorrect (Bilger et al., 1984). In our data the entire sentence was periodically interrupted and participant performance was scored across all key words in the sentence. Scoring only the last word may reduce individual differences in the ability to compensate across an entire sentence, because sentence context effects take place across an entire sentence (Stanovich and West, 1983; Kutas et al., 2019), the effect builds over time when sentences are predictable (Brothers and Kuperberg, 2021), and high predictability increases the benefit from glimpses of target speech across an entire sentence (Schoof and Rosen, 2015). However, it should be noted that the RSPIN High Silent, RSPIN High Noise, and RSPIN Low Noise sentences had a non-normal distribution of residuals in the current data set and thus these results should be interpreted with caution and RSPIN High Noise did not survive a Bonferroni correction for multiple comparisons.

The highest variability across listeners and poorest performance occurred for PRESTO sentences. This increase in variability and decrease in performance may have occurred for several reasons. First, the PRESTO sentences were designed to incorporate multiple factors during speech recognition: talker characteristics, dialect, and the role of higher order processes. The PRESTO sentences also vary in length and syntactic complexity. Taken together, these factors make PRESTO sentences less constrained than the RSPIN sentences and, thus, more representative of everyday speech (Cole et al., 2010). Second, the PRESTO sentences used here contain 455 unique words, which exceeds both the RSPIN High (421 words) and RSPIN Low (218 words) conditions. Therefore, the variability in the results may follow simply from the increased variability in the number of unique words in the PRESTO sentence set. Third, many of the key words in the PRESTO sentence set are longer and contain additional syllables (average of 7.96 syllables per sentence for PRESTO sentences compared to 6.14 syllables per sentence for the RSPIN High and 6.29 for the RSPIN Low context sentence sets). While

the silent gap and noise burst interruptions in the current experiment were designed to be shorter than the average syllabus nuclei duration in American English (Peterson and Lehiste, 1960), minimizing the obliteration of syllables entirely (Miller and Licklider, 1950), additional syllables in a key word may provide listeners with multiple glimpses at one word, which may improve or support perceptual restoration ability. This wider range of syllabic structure may contribute to the increased variability in the PRESTO sentences compared to the RSPIN sentences.

The role of working memory capacity in perceptual restoration

Working memory capacity measured using the Reading Span Task was significantly correlated with or acted as a significant predictor for interrupted speech recognition in five of the six sentence conditions. The significance of working memory capacity is in line with some previous literature for noise burst interrupted sentences using low-context, QuickSIN stimuli and a similar interruption paradigm, indicating the importance of working memory capacity for noise burst sentences (Millman and Mattys, 2017; Nagaraj and Magimairaj, 2017). However, previous literature has also found that working memory capacity does not play a role during perceptual restoration of PRESTO sentences using a very similar interruption process (Bologna et al., 2018). Bologna and colleagues used a zero signal-to-noise ratio (SNR) for the noise burst stimuli so that the speech was the same overall intensity as the noise bursts. This design may reduce the percept of speech continuity behind the noise burst, which may impede or interfere with the reprocessing role of working memory capacity during perceptual restoration, making the task more difficult than a noise burst condition with a negative SNR as in the current study. This would fall in line with existing literature that indicates few significant correlations between working memory capacity and silent gap interrupted sentences (Nagaraj and Knapp, 2015; Shafiro et al., 2015; Jaekel et al., 2018). A unique aspect of the current study is the wider range of participant age compared to most existing data, which tested only younger participants (Nagaraj and Knapp, 2015; Nagaraj and Magimairaj, 2017) or utilized group comparisons between older and younger adults, largely missing middle-aged listeners (Shafiro et al., 2015; Millman and Mattys, 2017; Bologna et al., 2018; Jaekel et al., 2018). Given the changing role that working memory capacity plays with increasing age (Wingfield et al., 1988), its effect on language comprehension (Caplan and Waters, 2005), and its possible task dependent nature (Turner and Engle, 1989), this may explain why the current data set found significant working memory capacity correlations for the majority of difficult silent gap conditions.

Lexical knowledge, lexical access speed, and perceptual restoration

The current data add to the evidence that lexical knowledge is important during perceptual restoration (Benard et al., 2014; Nagaraj and Magimairaj, 2017; Bologna et al., 2018; Jaekel et al., 2018). Under the ELU model, lexical knowledge is thought to support explicit working memory reprocessing within the episodic buffer (Rönnberg et al., 2013). This explicit reprocessing identifies likely and unlikely lexical candidates for the missing speech segments and attempts to reconcile segments into a cohesive, logical whole across the entire utterance (Bashford et al., 1992; Zhang and Samuel, 2018). In this way, the most lexically and contextually appropriate candidate can then be chosen by comparing options at the sentence level rather than just the gap level, thereby improving perceptual restoration across the entire utterance (Bashford and Warren, 1987).

For the silent gap interrupted conditions where lexical knowledge was strongly predictive, it is feasible that for listeners with greater lexical knowledge that a larger set of possible lexical candidates might be identified in the silent gap conditions than for listeners with poorer vocabularies. For the noise burst sentences where lexical knowledge was less predictive, it is possible that the noise burst itself may create enough illusory perceptual continuity that the correct lexical candidate can be more easily identified for all listeners, regardless of vocabulary size (Bashford et al., 1996). Alternatively, listeners with greater lexical knowledge may be less susceptible to misidentification of word boundaries in the silent gap conditions, making them better able to activate appropriate lexical candidates despite incomplete lexical neighborhood activation (Clarke et al., 2016). This alternative hypothesis follows the Neighborhood Activation Model, which suggests that listeners with greater lexical knowledge, even without priming, are better able to activate lexical neighborhoods with incomplete information, and that this effect is only detectable in the silent gap interrupted conditions because the noise burst sentences facilitate enough lexical neighborhood activation for all listeners (Luce and Pisoni, 1998; Luce et al., 2000).

To date, no known data have been reported on the relationship between lexical access speed, or the rate at which lexical candidates are identified and selected, and perceptual restoration. The current data do not support a significant role of lexical access speed. The lack of results for lexical access speed may stem from the LexTALE task itself, which was not designed to assess lexical access speed but rather to assess English language proficiency for English second language learners (Lemhöfer and Broersma, 2012), though it does have predictive value as a rapid task of proficiency assessment in English first language learners (Lorette and Dewaele, 2015). Although a computerized assessment does allow for the capture of reaction time for real words, non-words, and correct and incorrect items, the upper time limit of 2,000 ms may artificially limit lexical

access speed. Future studies of perceptual restoration and lexical access speed should include measures designed to capture this time-sensitive measure.

Inhibitory control, processing speed, and perceptual restoration

Inhibitory control was significantly correlated with and acted as a significant model predictor for three of the six sentence conditions: RSPIN Low Silent, RSPIN Low Noise, and PRESTO Noise. These results contrast those by [Bologna et al. \(2018\)](#), who found that inhibitory control did not significantly improve model fit for perceptual restoration in either silent gap or noise burst sentences. One possibility for the discrepancy between the current data and the results from [Bologna et al. \(2018\)](#) is the administration of the Stroop task. In the current data, the Stroop task was administered using an online platform, and listeners were asked to press a corresponding color key (e.g., “g” for green) when responding and do so as rapidly as possible. Remembering key location, key correspondence, and the motor control necessary to complete the task may have engaged working memory beyond what occurs during the process of responding verbally in the traditional administration of the Stroop task. Our Stroop task was correlated with and a significant predictor for most of the noise burst sentence conditions. This may indicate that inhibitory control plays an active role in inhibiting the irrelevant noise bursts when reprocessing speech fragments during perceptual restoration, and that listeners who are better able to inhibit the noise bursts are better able to focus on cognitive tasks that restore missing speech. However, RSPIN High Noise and RSPIN Low Noise conditions had a non-normal residuals distribution and this result should be interpreted with caution and RSPIN High Noise did not survive a Bonferroni correction for multiple comparisons.

Potential limitations of this experiment include the inability to measure audiometric thresholds from participants in this study, relying on self-report measures of “normal hearing.” Because auditory thresholds decline with increasing age ([Gates and Mills, 2005](#); [Huang and Tang, 2010](#)) it is possible that hearing acuity may have had an unmeasured impact on the results, despite the lack of significant correlations with both age and SSQ on cognitive/linguistic data and restoration of missing speech. Next, it should be noted that the cognitive/linguistic measures in this study were all in the visual modality while the outcome measures of interest were in the auditory modality. While many of the cognitive and linguistic measures included in this study are often thought of as domain-general (i.e., they are not modality specific), there is evidence that modality differences may affect how signals are processed cortically ([Salthouse and Meinz, 1995](#); [Crottaz-Herbette et al., 2004](#); [Roberts and Hall, 2008](#)) which may influence these results. Furthermore, the scoring method chosen for cognitive/linguistic measures can often yield different results and the results

from this study should be compared only to other measures administered in a similar fashion ([Knight and Heinrich, 2017](#)). Last, because the noise burst conditions are generally perceived as being less difficult than the silent gap conditions and the sentence types (e.g., RSPIN High, RSPIN Low, and PRESTO) differ from one another with regard to sentence and word length, these conditions may differ from one another with regard to overall task difficulty which can affect overall response accuracy ([Robinson, 2001](#)).

In the current experiment, processing speed, or the rate at which cognitive tasks are completed, was significantly correlated with RSPIN Low Silent, PRESTO Silent, and PRESTO Noise conditions and acted as a significant predictor in the RSPIN Low Silent and PRESTO Silent conditions. These results are similar to those found by [Bologna et al. \(2018\)](#) who found that interrupted key word recognition improved with faster processing speed when measured using the connections line making test ([Salthouse, 2000](#)). Given that processing speed was significant in two of three silent gap conditions, it is possible that these conditions are more difficult compared to the noise burst conditions and listeners who are able to reprocess and reanalyze the information more rapidly might be better able to restore missing speech information.

Conclusion

In this study, we hypothesized that higher-order cognitive and linguistic abilities would facilitate the restoration of missing speech information using the ELU model framework ([Rönnerberg et al., 2013](#)). The interrupted speech paradigm was utilized to explore this hypothesis, which in this case removed 50% of the speech signal in order to encourage participants to explicitly reprocess and reanalyze the incomplete speech signal. We predicted that listeners with stronger cognitive and linguistic abilities measured using validated cognitive measures would restore more missing speech information than those with weaker cognitive and linguistic abilities. Working memory capacity and lexical knowledge (i.e., vocabulary) played the most consistent and unique role in perceptual restoration across the sentence conditions, followed by inhibitory control and processing speed. In general, silent gap conditions appeared to be related to a broader range of cognitive and linguistic abilities whereas noise burst conditions were predicted by and correlated with working memory capacity and inhibitory control. Furthermore, sentences that had limited context cues and lacked predictability or were more like those encountered in everyday listening were significantly correlated with and predicted by a wider range of cognitive and linguistic abilities than those that contained additional context cues and had higher levels of predictability. The differences between silent gap and noise burst conditions as well as the context, predictability, and everyday speech conditions may be related to task-dependent difficulties that recruit different constellations

of cognitive and linguistic abilities to facilitate the restoration of missing speech information. In sum, perceptual restoration of speech is a complex process that relies on an individual's ability to store and reprocess, to identify potential lexical candidates, to inhibit irrelevant information, to contextually consider several options simultaneously, and to complete these cognitive tasks rapidly, and listeners vary considerably in these abilities (Carroll and Maxwell, 1979; Reuter-Lorenz et al., 2000; Cabeza et al., 2002; George et al., 2007; Rudner et al., 2008; Boogert et al., 2018).

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by Northwestern University Institutional Review Board. The patients/participants provided their written informed consent to participate in this study.

Author contributions

Both authors contributed to study design, data collection, management, analysis, and manuscript preparation and approved the submitted version.

References

- 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. doi: 10.1080/14992020802301142
- Bashford, J. A., and Warren, R. M. (1987). Multiple phonemic restorations follow the rules for auditory induction. *Percept. Psychophys.* 42, 114–121. doi: 10.3758/BF03210499
- Bashford, J. A., Riener, K. R., and Warren, R. M. (1992). Increasing the intelligibility of speech through multiple phonemic restorations. *Percept. Psychophys.* 51, 211–217. doi: 10.3758/BF03212247
- Bashford, J. A., Warren, R. M., and Brown, C. A. (1996). Use of speech-modulated noise adds strong “bottom-up” cues for phonemic restoration. *Percept. Psychophys.* 58, 342–350. doi: 10.3758/BF03206810
- Başkent, D. (2010). Phonemic restoration in sensorineural hearing loss does not depend on baseline speech perception scores. *J. Acoust. Soc. Am.* 128, EL169–EL174. doi: 10.1121/1.3475794
- Başkent, D., Clarke, J., Pals, C., Benard, M. R., Bhargava, P., Saija, J., et al. (2016). Cognitive compensation of speech perception with hearing impairment, cochlear implants, and aging: How and to what degree can it be achieved? *Trends Hear.* 20:2331216516670279. doi: 10.1177/2331216516670279
- Başkent, D., Eiler, C., and Edwards, B. (2009). Effects of envelope discontinuities on perceptual restoration of amplitude-compressed speech. *J. Acoust. Soc. Am.* 125, 3995–4005. doi: 10.1121/1.3125329
- Benard, M. R., and Başkent, D. (2014). Perceptual learning of temporally interrupted spectrally degraded speech. *J. Acoust. Soc. Am.* 136, 1344–1351. doi: 10.1121/1.4892756
- Benard, M. R., Mensink, J. S., and Başkent, D. (2014). Individual differences in top-down restoration of interrupted speech: Links to linguistic and cognitive abilities. *J. Acoust. Soc. Am.* 135, EL88–EL94. doi: 10.1121/1.4862879
- Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. Series B Methodol.* 57, 289–300. doi: 10.1111/j.2517-6161.1995.tb02031.x
- Bilger, R. C., Nuetzel, J. M., Rabinowitz, W. M., and Rzeczkowski, C. (1984). Standardization of a test of speech perception in noise. *J. Speech Lang. Hear. Res.* 27, 32–48. doi: 10.1044/jshr.2701.32
- Bologna, W. J., Vaden, K. I. Jr., Ahlstrom, J. B., and Dubno, J. R. (2018). Age effects on perceptual organization of speech: Contributions of glimpsing, phonemic restoration, and speech segregation. *J. Acoust. Soc. Am.* 144, 267–281. doi: 10.1121/1.5044397
- Bologna, W. J., Vaden, K. I. Jr., Ahlstrom, J. B., and Dubno, J. R. (2019). Age effects on the contributions of envelope and periodicity cues to recognition of interrupted speech in quiet and with a competing talker. *J. Acoust. Soc. Am.* 145, EL173–EL178. doi: 10.1121/1.5091664
- Boogert, N. J., Madden, J. R., Morand-Ferron, J., and Thornton, A. (2018). Measuring and understanding individual differences in cognition. *R. Soc.* 373, 1–10. doi: 10.1098/rstb.2017.0280

Funding

This work was partially supported by NIH (R01 DC006014).

Acknowledgments

The authors would like to acknowledge the other members of the Hearing Aid Laboratory who offered their guidance, support, and advice during the completion of this project.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Borrie, S. A., Barrett, T. S., and Yoho, S. E. (2019). Autoscore: An open-source automated tool for scoring listener perception of speech. *J. Acoust. Soc. Am.* 145, 392–399. doi: 10.1121/1.5087276
- Breusch, T. S., and Pagan, A. R. (1979). A simple test for heteroscedasticity and random coefficient variation. *Econometrica* 47, 1287–1294. doi: 10.2307/1911963
- Brookshire, R. H., and Nicholas, L. (1984). Comprehension of directly and indirectly stated main ideas and details in discourse by brain-damaged and non-brain-damaged listeners. *Brain Lang.* 21, 21–36. doi: 10.1016/0093-934X(84)90033-6
- Brothers, T., and Kuperberg, G. R. (2021). Word predictability effects are linear, not logarithmic: Implications for probabilistic models of sentence comprehension. *J. Mem. Lang.* 116:104174. doi: 10.1016/j.jml.2020.104174
- Bursac, Z., Gauss, C. H., Williams, D. K., and Hosmer, D. W. (2008). Purposeful selection of variables in logistic regression. *Source Code Biol. Med.* 3, 1–8. doi: 10.1186/1751-0473-3-17
- Cabeza, R., Anderson, N. D., Locantore, J. K., and Mcintosh, A. R. (2002). Aging gracefully: Compensatory brain activity in high-performing older adults. *Neuroimage* 17, 1394–1402. doi: 10.1006/nimg.2002.1280
- Caplan, D., and Waters, G. (2005). The relationship between age, processing speed, working memory capacity, and language comprehension. *Memory* 13, 403–413. doi: 10.1080/09658210344000459
- Carroll, J. B., and Maxwell, S. E. (1979). Individual differences in cognitive abilities. *Annu. Rev. Psychol.* 30, 603–640. doi: 10.1146/annurev.ps.30.020179.003131
- Chatterjee, M., Peredo, F., Nelson, D., and Başkent, D. (2010). Recognition of interrupted sentences under conditions of spectral degradation. *J. Acoust. Soc. Am.* 127, EL37–EL41. doi: 10.1121/1.3284544
- Clarke, J., Başkent, D., and Gaudrain, E. (2016). Pitch and spectral resolution: A systematic comparison of bottom-up cues for top-down repair of degraded speech. *J. Acoust. Soc. Am.* 139, 395–405. doi: 10.1121/1.4939962
- Cole, J., Mo, Y., and Baek, S. (2010). The role of syntactic structure in guiding prosody perception with ordinary listeners and everyday speech. *Lang. Cogn. Process.* 25, 1141–1177. doi: 10.1080/01690960903525507
- Cooke, M. (2006). A glimpsing model of speech perception in noise. *J. Acoust. Soc. Am.* 119, 1562–1573. doi: 10.1121/1.2166600
- Crottaz-Herbette, S., Anagnoson, R. T., and Menon, V. (2004). Modality effects in verbal working memory: Differential prefrontal and parietal responses to auditory and visual stimuli. *Neuroimage* 21, 340–351. doi: 10.1016/j.neuroimage.2003.09.019
- D'Agostino, R. B. (1971). An omnibus test of normality for moderate and large size samples. *Biometrika* 58, 341–348. doi: 10.1093/biomet/58.2.341
- Daneman, M., and Carpenter, P. A. (1980). Individual differences in working memory and reading. *J. Verbal Learn. Verbal Behav.* 19, 450–466. doi: 10.1016/S0022-5371(80)90312-6
- Daneman, M., and Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychon. Bull. Rev.* 3, 422–433. doi: 10.3758/BF03214546
- Dey, A., and Sommers, M. S. (2015). Age-related differences in inhibitory control predict audiovisual speech perception. *Psychol. Aging* 30:634. doi: 10.1037/pag0000033
- Dijkgraaf, A., Hartsuiker, R. J., and Duyck, W. (2016). Predicting upcoming information in native-language and non-native-language auditory word recognition. *Bilingualism Lang. Cogn.* 20, 917–930. doi: 10.1017/S1366728916000547
- Dryden, A., Allen, H. A., Henshaw, H., and Heinrich, A. (2017). The association between cognitive performance and speech-in-noise perception for adult listeners: A systematic literature review and meta-analysis. *Trends Hear.* 21:2331216517744675. doi: 10.1177/2331216517744675
- Ellis, R. J., Molander, P., Rönnberg, J., Lyxell, B., Andersson, G., and Lunner, T. (2016). Predicting speech-in-noise recognition from performance on the trail making test: Results from a large-scale internet study. *Ear Hear.* 37, 73–79. doi: 10.1097/AUD.0000000000000218
- Eriksen, B. A., and Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Percept. Psychophys.* 16, 143–149. doi: 10.3758/BF03203267
- Fogerty, D., Ahlstrom, J. B., Bologna, W. J., and Dubno, J. R. (2015). Sentence intelligibility during segmental interruption and masking by speech-modulated noise: Effects of age and hearing loss. *J. Acoust. Soc. Am.* 137, 3487–3501. doi: 10.1121/1.4921603
- Friedman, N. P., and Miyake, A. (2005). Comparison of four scoring methods for the reading span test. *Behav. Res. Methods* 37, 581–590. doi: 10.3758/BF03192728
- Füllgrabe, C., Moore, B. C., and Stone, M. A. (2015). Age-group differences in speech identification despite matched audiometrically normal hearing: Contributions from auditory temporal processing and cognition. *Front. Aging Neurosci.* 6:347. doi: 10.3389/fnagi.2014.00347
- Gatehouse, S., and Noble, W. (2004). The speech, spatial and qualities of hearing scale (SSQ). *Int. J. Audiol.* 43, 85–99. doi: 10.1080/14992020400050014
- Gates, G. A., and Mills, J. H. (2005). Presbycusis. *Lancet* 366, 1111–1120. doi: 10.1016/S0140-6736(05)67423-5
- George, E. L., Zekveld, A. A., Kramer, S. E., Goverts, S. T., Festen, J. M., and Houtgast, T. (2007). Auditory and nonauditory factors affecting speech reception in noise by older listeners. *J. Acoust. Soc. Am.* 121, 2362–2375. doi: 10.1121/1.2642072
- Golden, C. J. (1976). Identification of brain disorders by the stroop color and word test. *J. Clin. Psychol.* 32, 654–658. doi: 10.1002/1097-4679(197607)32:3<654::AID-JCLP2270320336>3.0.CO;2-Z
- Gordon-Salant, S., and Fitzgibbons, P. J. (2004). Effects of stimulus and noise rate variability on speech perception by younger and older adults. *J. Acoust. Soc. Am.* 115, 1808–1817. doi: 10.1121/1.1645249
- Harris, P. A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., and Conde, J. G. (2009). Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *J. Biomed. Inform.* 42, 377–381. doi: 10.1016/j.jbi.2008.08.010
- Hilbert, S., Nakagawa, T. T., Puci, P., Zech, A., and Bühner, M. (2015). The digit span backwards task: Verbal and visual cognitive strategies in working memory assessment. *Eur. J. Psychol. Assess.* 31, 174–180. doi: 10.1027/1015-5759/a000223
- Hosmer, D. W. Jr., Lemeshow, S., and Sturdivant, R. X. (2013). *Applied logistic regression*. Hoboken, NJ: John Wiley & Sons. doi: 10.1002/9781118548387
- Huang, Q., and Tang, J. (2010). Age-related hearing loss or presbycusis. *Eur. Arch. Otorhinolaryngol.* 267, 1179–1191. doi: 10.1007/s00405-010-1270-7
- Jaekel, B. N., Newman, R. S., and Goupell, M. J. (2018). Age effects on perceptual restoration of degraded interrupted sentences. *J. Acoust. Soc. Am.* 143, 84–97. doi: 10.1121/1.5016968
- Jensen, A. R. (1965). Scoring the Stroop test. *Acta Psychol.* 24, 398–408. doi: 10.1016/0001-6918(65)90024-7
- Jenstad, L. M., and Souza, P. E. (2007). Temporal envelope changes of compression and speech rate: Combined effects on recognition for older adults. *J. Speech Lang. Hear. Res.* 50, 1123–1138. doi: 10.1044/1092-4388(2007)078
- Jin, S.-H., and Nelson, P. B. (2010). Interrupted speech perception: The effects of hearing sensitivity and frequency resolution. *J. Acoust. Soc. Am.* 128, 881–889. doi: 10.1121/1.3458851
- Johnston, R., Jones, K., and Manley, D. (2018). Confounding and collinearity in regression analysis: A cautionary tale and an alternative procedure, illustrated by studies of British voting behaviour. *Qual. Quant.* 52, 1957–1976. doi: 10.1007/s11135-017-0584-6
- Jones, G., and Macken, B. (2015). Questioning short-term memory and its measurement: Why digit span measures long-term associative learning. *Cognition* 144, 1–13. doi: 10.1016/j.cognition.2015.07.009
- Kassambara, A. (2021). *rstatix: Pipe-friendly framework for basic statistical tests. R package version 0.7.0 ed.* <https://CRAN.R-project.org/package=rstatix>
- Keuleers, E., Stevens, M., Mandera, P., and Brysbaert, M. (2015). Word knowledge in the crowd: Measuring vocabulary size and word prevalence in a massive online experiment. *Q. J. Exp. Psychol.* 68, 1665–1692. doi: 10.1080/17470218.2015.1022560
- Kidd, G. R., and Humes, L. E. (2012). Effects of age and hearing loss on the recognition of interrupted words in isolation and in sentences. *J. Acoust. Soc. Am.* 131, 1434–1448. doi: 10.1121/1.3675975
- Kline, R. B. (2015). *Principles and practice of structural equation modeling*. New York, NY: Guilford publications.
- Knight, S., and Heinrich, A. (2017). Different measures of auditory and visual stroop interference and their relationship to speech intelligibility in noise. *Front. Psychol.* 8:230. doi: 10.3389/fpsyg.2017.00230
- Kutas, M., Lindamood, T. E., and Hillyard, S. A. (2019). “Word expectancy and event-related brain potentials during sentence processing” in *Preparatory states & processes*, eds S. Kornblum and J. Requin (London: Psychology Press). doi: 10.4324/9781315792385-11
- Langenecker, S. A., Nielson, K. A., and Rao, S. M. (2004). fMRI of healthy older adults during stroop interference. *Neuroimage* 21, 192–200. doi: 10.1016/j.neuroimage.2003.08.027

- Lansbergen, M. M., Kenemans, J. L., and Van Engeland, H. (2007). Stroop interference and attention-deficit/hyperactivity disorder: A review and meta-analysis. *Neuropsychology* 21:251. doi: 10.1037/0894-4105.21.2.251
- Lehto, J. (1996). Are executive function tests dependent on working memory capacity? *Q. J. Exp. Psychol. A* 49, 29–50. doi: 10.1080/713755616
- Lemhöfer, K., and Broersma, M. (2012). Introducing lexTALE: A quick and valid lexical test for advanced learners of English. *Behav. Res. Methods* 44, 325–343. doi: 10.3758/s13428-011-0146-0
- Levene, H. (ed.) (1960). *Robust tests for equality of variances*. Stanford, CA: Stanford University Press.
- Li, N., and Loizou, P. C. (2007). Factors influencing glimpsing of speech in noise. *J. Acoust. Soc. Am.* 122, 1165–1172. doi: 10.1121/1.2749454
- Lorette, P., and Dewaele, J.-M. (2015). Emotion recognition ability in English among L1 and L2 users of English. *Int. J. Lang. Cult.* 2, 62–86. doi: 10.1075/ijolc.2.1.03lor
- Luce, P. A., and Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear Hear.* 19:1. doi: 10.1097/00003446-199802000-00001
- Luce, P. A., Goldinger, S. D., Auer, E. T., and Vitevitch, M. S. (2000). Phonetic priming, neighborhood activation, and PARSYN. *Percept. Psychophys.* 62, 615–625. doi: 10.3758/BF03212113
- Mansfield, E. R., and Helms, B. P. (1982). Detecting multicollinearity. *Am. Statistician* 36, 158–160. doi: 10.1080/00031305.1982.10482818
- Marslen-Wilson, W. D., and Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cogn. Psychol.* 10, 29–63. doi: 10.1016/0010-0285(78)90018-X
- Mickey, R. M., and Greenland, S. (1989). The impact of confounder selection criteria on effect estimation. *Am. J. Epidemiol.* 129, 125–137. doi: 10.1093/oxfordjournals.aje.a115101
- Miller, G. A., and Licklider, J. C. (1950). The intelligibility of interrupted speech. *J. Acoust. Soc. Am.* 22, 167–173. doi: 10.1121/1.1906584
- Millman, R. E., and Mattys, S. L. (2017). Auditory verbal working memory as a predictor of speech perception in modulated maskers in listeners with normal hearing. *J. Speech Lang. Hear. Res.* 60, 1236–1245. doi: 10.1044/2017_JSLHR-S-16-0105
- Moder, K. (2007). How to keep the type I error rate in ANOVA if variances are heteroscedastic. *Austrian J. Stat.* 36, 179–188. doi: 10.17713/ajs.v36i3.329
- Morrison, C. M., and Gibbons, Z. C. (2006). Lexical determinants of semantic processing speed. *Vis. Cogn.* 13, 949–967. doi: 10.1080/13506280544000129
- Nagaraj, N. K., and Knapp, A. N. (2015). No evidence of relation between working memory and perception of interrupted speech in young adults. *J. Acoust. Soc. Am.* 138, EL145–EL150. doi: 10.1121/1.4927635
- Nagaraj, N. K., and Magimairaj, B. M. (2017). Role of working memory and lexical knowledge in perceptual restoration of interrupted speech. *J. Acoust. Soc. Am.* 142, 3756–3766. doi: 10.1121/1.5018429
- Neuman, A. C., Wroblewski, M., Hajicek, J., and Rubinstein, A. (2010). Combined effects of noise and reverberation on speech recognition performance of normal-hearing children and adults. *Ear Hear.* 31, 336–344. doi: 10.1097/AUD.0b013e3181d3d514
- Oxenham, A. J. (2008). Pitch perception and auditory stream segregation: Implications for hearing loss and cochlear implants. *Trends Amplif.* 12, 316–331. doi: 10.1177/1084713808325881
- Paap, K. R., Anders-Jefferson, R., Zimiga, B., Mason, L., and Mikulinsky, R. (2020). Interference scores have inadequate concurrent and convergent validity: Should we stop using the flanker, Simon, and spatial Stroop tasks? *Cogn. Res. Princ. Implic.* 5, 1–27. doi: 10.1186/s41235-020-0207-y
- Perron, M., Dimitrijevic, A., and Alain, C. (2022). Objective and subjective hearing difficulties are associated with lower inhibitory control. *Ear Hear.* 43, 1904–1916. doi: 10.1097/AUD.0000000000001227
- Peterson, G. E., and Lehiste, I. (1960). Duration of syllable nuclei in English. *J. Acoust. Soc. Am.* 32, 693–703. doi: 10.1121/1.1908183
- Powers, G. L., and Wilcox, J. C. (1977). Intelligibility of temporally interrupted speech with and without intervening noise. *J. Acoust. Soc. Am.* 61, 195–199. doi: 10.1121/1.381255
- R Core Team (2013). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.
- Reuter-Lorenz, P. A., Jonides, J., Smith, E. E., Hartley, A., Miller, A., Marshuetz, C., et al. (2000). Age differences in the frontal lateralization of verbal and spatial working memory revealed by PET. *J. Cogn. Neurosci.* 12, 174–187. doi: 10.1162/089892900561814
- Roberts, K. L., and Hall, D. A. (2008). Examining a supramodal network for conflict processing: A systematic review and novel functional magnetic resonance imaging data for related visual and auditory Stroop tasks. *J. Cogn. Neurosci.* 20, 1063–1078. doi: 10.1162/jocn.2008.20074
- Robinson, P. (2001). Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Appl. Linguist.* 22, 27–57. doi: 10.1093/applin/22.1.27
- Rönnerberg, J., Arlinger, S., Lyxell, B. R., and Kinnefors, C. (1989). Visual evoked potentials: Relation to adult speechreading and cognitive function. *J. Speech Lang. Hear. Res.* 32, 725–735. doi: 10.1044/jshr.3204.725
- Rönnerberg, J., Lunner, T., Zekveld, A., Sörqvist, P., Danielsson, H., Lyxell, B., et al. (2013). The ease of language understanding (ELU) model: Theoretical, empirical, and clinical advances. *Front. Syst. Neurosci.* 7:31. doi: 10.3389/fnsys.2013.00031
- Rozas, A. X. P., Juncos-Rabadán, O., and González, M. S. R. (2008). Processing speed, inhibitory control, and working memory: Three important factors to account for age-related cognitive decline. *Int. J. Aging Hum. Dev.* 66, 115–130. doi: 10.2190/AG.66.2.b
- Rudner, M., Foo, C., Sundewall-Thoren, E., Lunner, T., and Rönnerberg, J. (2008). Phonological mismatch and explicit cognitive processing in a sample of 102 hearing-aid users. *Int. J. Audiol.* 47, S91–S98. doi: 10.1080/14992020802304393
- Salthouse, T. A. (1992). Influence of processing speed on adult age differences in working memory. *Acta Psychol.* 79, 155–170. doi: 10.1016/0001-6918(92)90030-H
- Salthouse, T. A. (2000). Aging and measures of processing speed. *Biol. Psychol.* 54, 35–54. doi: 10.1016/S0301-0511(00)00052-1
- Salthouse, T. A., and Meinz, E. J. (1995). Aging, inhibition, working memory, and speed. *J. Gerontol. B Psychol. Sci. Soc. Sci.* 50, 297–306. doi: 10.1093/geronb/50B.6.P297
- Sanders, L. M., Hortobágyi, T., Balasingham, M., Van Der Zee, E. A., and Van Heuvelen, M. J. (2018). Psychometric properties of a flanker task in a sample of patients with dementia: A pilot study. *Dement. Geriatr. Cogn. Dis. Extra* 8, 382–392. doi: 10.1159/000493750
- Scarpina, F., and Tagini, S. (2017). The Stroop color and word test. *Front. Psychol.* 8:557. doi: 10.3389/fpsyg.2017.00557
- Schoof, T., and Rosen, S. (2015). High sentence predictability increases the fluctuating masker benefit. *J. Acoust. Soc. Am.* 138, EL181–EL186. doi: 10.1121/1.4929627
- Shafiro, V., Sheft, S., and Risley, R. (2016). The intelligibility of interrupted and temporally altered speech: Effects of context, age, and hearing loss. *J. Acoust. Soc. Am.* 139, 455–465. doi: 10.1121/1.4939891
- Shafiro, V., Sheft, S., Risley, R., and Gygi, B. (2015). Effects of age and hearing loss on the intelligibility of interrupted speech. *J. Acoust. Soc. Am.* 137, 745–756. doi: 10.1121/1.4906275
- Smith, S., Krizman, J., Liu, C., White-Schwoch, T., Nicol, T., and Kraus, N. (2019). Investigating peripheral sources of speech-in-noise variability in listeners with normal audiograms. *Hear. Res.* 371, 66–74. doi: 10.1016/j.heares.2018.11.008
- Stanovich, K. E., and West, R. F. (1983). On priming by a sentence context. *J. Exp. Psychol.* 112, 1–36. doi: 10.1037/0096-3445.112.1.1
- Stenbäck, V., Marsja, E., Hällgren, M., Lyxell, B., and Larsby, B. (2021). The contribution of age, working memory capacity, and inhibitory control on speech recognition in noise in young and older adult listeners. *J. Speech Lang. Hear. Res.* 64, 4513–4523. doi: 10.1044/2021_JSLHR-20-00251
- Tamati, T. N., Gilbert, J. L., and Pisoni, D. B. (2013). Some factors underlying individual differences in speech recognition on PRESTO: A first report. *J. Am. Acad. Audiol.* 24, 616–634. doi: 10.3766/jaaa.24.7.10
- Thompson, C. G., Kim, R. S., Aloe, A. M., and Becker, B. J. (2017). Extracting the variance inflation factor and other multicollinearity diagnostics from typical regression results. *Basic Appl. Soc. Psychol.* 39, 81–90. doi: 10.1080/01973533.2016.1277529
- Turner, M. L., and Engle, R. W. (1989). Is working memory capacity task dependent? *J. Mem. Lang.* 28, 127–154. doi: 10.1016/0749-596X(89)90040-5
- Turner, M., and Ridsdale, J. (2004). *The digit memory test*. Baltimore, MD: Dyslexia International.
- Van der Elst, W., Van Boxtel, M. P. J., Van Breukelen, G. J. P., and Jolles, J. (2006). The Stroop color-word test: Influence of age, sex, and education; and normative data for a large sample across the adult age range. *Assessment* 13, 62–79. doi: 10.1177/1073191105283427
- Venables, W., and Ripley, B. (2002). *Modern applied statistics with S*, 4th Edn. New York, NY: Springer. doi: 10.1007/978-0-387-21706-2
- Verhaeghen, P. (2003). Aging and vocabulary score: A meta-analysis. *Psychol. Aging* 18:332. doi: 10.1037/0882-7974.18.2.332

- Vittinghoff, E., Glidden, D. V., Shiboski, S. C., and McCulloch, C. E. (2006). *Regression methods in biostatistics: Linear, logistic, survival, and repeated measures models*. Berlin: Springer Science & Business Media.
- Warren, R. M. (1970). Perceptual restoration of missing speech sounds. *Science* 167, 392–393. doi: 10.1126/science.167.3917.392
- Warren, R. M., and Obusek, C. J. (1971). Speech perception and phonemic restorations. *Percept. Psychophys.* 9, 358–362. doi: 10.3758/BF03212667
- Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. New York, NY: Springer-Verlag. doi: 10.1007/978-3-319-24277-4
- Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D. A., François, R., et al. (2019). Welcome to the tidyverse. *J. Open Source Softw.* 4:1686. doi: 10.21105/joss.01686
- Wickham, H., François, R., Henry, L., and Müller, K. (2022). *dplyr: A grammar of data manipulation*. <https://CRAN.R-project.org/package=dplyr> (accessed October 9, 2022).
- Wingfield, A., Stine, E. A., Lahar, C. J., and Aberdeen, J. S. (1988). Does the capacity of working memory change with age? *Exp. Aging Res.* 14, 103–107. doi: 10.1080/03610738808259731
- Woods, K. J., Siegel, M. H., Traer, J., and McDermott, J. H. (2017). Headphone screening to facilitate web-based auditory experiments. *Atten. Percept. Psychophys.* 79, 2064–2072. doi: 10.3758/s13414-017-1361-2
- Yumba, W. K. (2017). Cognitive processing speed, working memory, and the intelligibility of hearing aid-processed speech in persons with hearing impairment. *Front. Psychol.* 8:1308. doi: 10.3389/fpsyg.2017.01308
- Zeileis, A., and Hothorn, T. (2002). Diagnostic checking in regression relationships. *R News* 2, 7–10.
- Zhang, X., and Samuel, A. G. (2018). Is speech recognition automatic? Lexical competition, but not initial lexical access, requires cognitive resources. *J. Mem. Lang.* 100, 32–50. doi: 10.1016/j.jml.2018.01.002
- Zhang, Z. (2016). Model building strategy for logistic regression: Purposeful selection. *Ann. Transl. Med.* 4:111. doi: 10.21037/atm.2016.02.15