



One Size Does Not Fit All: Examining the Effects of Working Memory Capacity on Spoken Word Recognition in Older Adults Using Eye Tracking

OPEN ACCESS

Edited by:

Leah Fostick,
Ariel University, Israel

Reviewed by:

Subong Kim,
Purdue University, United States
Likun Zhan,
Beijing Language and Culture
University, China

*Correspondence:

Boaz M. Ben-David
boaz.ben.david@idc.ac.il

†ORCID:

Gal Nitsan
orcid.org/0000-0001-8567-2823
Karen Banai
orcid.org/0000-0002-2990-0470
Boaz M. Ben-David
orcid.org/0000-0002-0392-962X

Specialty section:

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

Received: 22 December 2021

Accepted: 14 March 2022

Published: 11 April 2022

Citation:

Nitsan G, Banai K and
Ben-David BM (2022) One Size Does
Not Fit All: Examining the Effects
of Working Memory Capacity on
Spoken Word Recognition in Older
Adults Using Eye Tracking.
Front. Psychol. 13:841466.
doi: 10.3389/fpsyg.2022.841466

Gal Nitsan^{1,2†}, Karen Banai^{2†} and Boaz M. Ben-David^{1,3,4*†}

¹ Baruch Ivcher School of Psychology, Reichman University (IDC), Herzliya, Israel, ² Department of Communication Sciences and Disorders, University of Haifa, Haifa, Israel, ³ Department of Speech-Language Pathology, University of Toronto, Toronto, ON, Canada, ⁴ Toronto Rehabilitation Institute, University Health Networks, Toronto, ON, Canada

Difficulties understanding speech form one of the most prevalent complaints among older adults. Successful speech perception depends on top-down linguistic and cognitive processes that interact with the bottom-up sensory processing of the incoming acoustic information. The relative roles of these processes in age-related difficulties in speech perception, especially when listening conditions are not ideal, are still unclear. In the current study, we asked whether older adults with a larger working memory capacity process speech more efficiently than peers with lower capacity when speech is presented in noise, with another task performed in tandem. Using the Eye-tracking of Word Identification in Noise Under Memory Increased Load (E-WINDMIL) an adapted version of the “visual world” paradigm, 36 older listeners were asked to follow spoken instructions presented in background noise, while retaining digits for later recall under low (single-digit) or high (four-digits) memory load. In critical trials, instructions (e.g., “point at the candle”) directed listeners’ gaze to pictures of objects whose names shared onset or offset sounds with the name of a competitor that was displayed on the screen at the same time (e.g., candy or sandal). We compared listeners with different memory capacities on the time course for spoken word recognition under the two memory loads by testing eye-fixations on a named object, relative to fixations on an object whose name shared phonology with the named object. Results indicated two trends. (1) For older adults with lower working memory capacity, increased memory load did not affect online speech processing, however, it impaired offline word recognition accuracy. (2) The reverse pattern was observed for older adults with higher working memory capacity: increased task difficulty significantly decreases online speech processing efficiency but had no effect on offline word recognition accuracy. Results suggest that in older adults, adaptation to adverse listening conditions is at least partially supported by cognitive

reserve. Therefore, additional cognitive capacity may lead to greater resilience of older listeners to adverse listening conditions. The differential effects documented by eye movements and accuracy highlight the importance of using both online and offline measures of speech processing to explore age-related changes in speech perception.

Keywords: speech perception, working memory, aging, word recognition, eye-tracking, visual world paradigm, cognitive hearing science

INTRODUCTION

A recent report by the World Health Organization (2021) emphasizes the importance of functional ability as a key to healthy aging. It suggests that preserving the abilities to build and maintain relationships and to grow learn and make decisions all promote well-being and healthy aging. These functional abilities depend heavily on successful speech perception. Indeed, difficulties understanding speech are one of the most prevalent complaints among older adults, especially in daily listening situations when listening conditions are not ideal (e.g., Abrams and Farrell, 2011). Although hearing deficits are a main source of difficulty in speech perception (Humes et al., 1994; Humes, 2021), successful speech perception also depends on the interaction of bottom-up hearing related factors and top-down linguistic and cognitive processes (Sommers, 2005; Zekveld et al., 2006; Pichora-Fuller, 2008; Rogers and Peelle, 2021). Furthermore, difficulties in speech perception are also observed among older adults with relatively preserved hearing (Sommers and Danielson, 1999; Fostick et al., 2013; Lash et al., 2013). Our goal is to test whether older listeners with a higher working memory capacity process speech in adverse conditions more efficiently than peers with lower capacity.

Previous studies in cognitive hearing science reported an association between individual differences in cognitive factors and differences in speech perception, even in young and healthy hearing populations. One consistent finding is that these differences are pronounced mainly when using complex testing materials (i.e., sentences, connected discourse comprehension, conversational situations; e.g., Heinrich et al., 2015; Dryden et al., 2017; Meister, 2017). For example, by comparing performance of older listeners across a wide range of speech perception tests differing in complexity, Heinrich et al. (2015) showed that the contribution of cognition increases as the complexity of the speech perception task increases. That is, for older adults, cognitive factors predict sentence perception to a larger extent than single spoken word perception. Of the many cognitive constructs tested, working memory has been widely recognized as related to differences in speech perception abilities, especially in adverse listening condition for older adults (see Akeroyd, 2008; Besser et al., 2013; Dryden et al., 2017 for relevant reviews). In particular, the storage and processing components of working memory play an important role in sentence processing as the listener is required to correctly encode the speech sounds, identify them as words, and then retain the string of words in memory until the sentence is fully heard (Daneman and Carpenter, 1980; Pichora-Fuller et al., 1995; Daneman and Merikle, 1996; Rönnberg et al., 2008). Working memory has also been linked

with inhibition of irrelevant information (Awh and Vogel, 2008). The latter is directly related to successful speech perception, where the listener needs to continuously inhibit irrelevant lexical items from his/her mental lexicon to allow correct word recognition. For example, Janse (2012) showed that when speech is presented in background noise, poor inhibitory abilities lead to greater interference by the competing noise which impairs speech perception of older adults.

Contrary to the agreement regarding the association between working memory, aging and spoken sentence processing, only little and mixed evidence is available on this association at the single word level. This is of special importance because lexical ambiguities frequently occur in daily life. For example, cell phones may distort a critical portion of the incoming signal. Consider the sentence “Grandpa! Have you seen the *dog*?” The word *dog* may be mistaken for *doll* (as the two share onset sounds, e.g., see Allopenna et al., 1998; Onset Cohort model, Marslen-Wilson, 1990; Shortlist, Norris, 1994) which can lead to miscommunication with severe consequences on future social participation. Despite these challenges, listeners appear to recognize words with little effort. Moreover, studying the effects of working memory at the single word level has theoretical implications. As spoken sentence processing involves many intervening factors, they may inflate the effects of working memory. Among the abilities necessary to understand sentences are sustained attention for the duration of the sentence and maintaining a running memory of the input to relate what is being heard to what has just been heard and to integrate it with what is about to be heard (Ayasse et al., 2017; Harel-Arbeli et al., 2021). Further, spoken context processing may be more influenced by linguistic experience and vocabulary than the processing of a single spoken word (Stine-Morrow et al., 2006; Borovsky et al., 2012; Ben-David et al., 2015; Kavé and Halamish, 2015). Thus, the aforementioned effects of working memory on the sentence level may reflect other processes. However, if effects are found at a single word level, that would indicate that working memory is involved at very early and basic levels of lexical access.

There is mixed evidence in the literature with regards to the effects of working memory and single spoken word recognition in aging. For example, Heinrich and Knight (2016) found that older adults’ performance in a visual working memory task significantly correlated with their performance on a word in noise (WIN) recognition task, irrespective of the noise level [both in low and high signal to noise ratios (SNRs)]. Gordon-Salant and Cole (2016) found similar results with both young and older adults, correlating auditory working memory capacity, with single-word recognition in noise. Conversely, other studies failed to find this correlation on the single word level.

For example, Parbery-Clark et al. (2011) did not find auditory working memory performance to correlate with performance on the WIN test for older adults with and without hearing loss. Similar findings were reported by Smith and Pichora-Fuller (2015) who failed to find a correlation between auditory and visual working memory performance and scores on the WIN test.

A possible explanation for these contradictory findings may stem from the use of offline measures to gauge word recognition (such as accuracy or SNR to achieve 50% recognition). Offline measures test the result of successful (or unsuccessful) word recognition, after the entire word has been heard, processed and a response has been made. It gages the final outcome of the process, and it cannot reveal the early processes underlying online speech processing. Additionally, previous works showed that this association between working memory and word recognition might differ depending on whether verbal or non-verbal measures of working memory are used and the modality of working memory tasks: auditory or visual. There is some evidence to suggest that auditory working memory plays a greater role in speech perception than visual working memory (Baldwin and Ash, 2011; Smith and Pichora-Fuller, 2015; Smith et al., 2016; Kim et al., 2020). Finally, none of the studies listed above tapped cognitive resources while performing speech recognition task, they only measured the correlation between performance on these separate measures. Direct manipulations of the memory load can allow us to better assess the causal relationship between reduced cognitive capacity and spoken word processing in aging.

The Current Study

In the current study, we examined the role of working memory capacity in spoken word recognition in adverse conditions for older adults. We hypothesized that older listeners with a larger working memory capacity would process speech more efficiently than their peers with a lower capacity; this is tested when speech is presented in noise, with another working memory demanding task performed in tandem. As listeners with lower working memory capacity already have fewer cognitive resources, we expect that the effects of increased load would be especially detrimental for their spoken word processing. This was tested using an adapted version of the eye-tracking “visual world” paradigm, coined the Eye-tracking of Word Identification in Noise Under Memory Increased Load (E-WINDMIL; Hadar et al., 2016; Nitsan et al., 2019). This paradigm was found to have significant test retest reliability for older adults (Baharav et al., 2021). In the E-WINDMIL listeners are instructed to press on one of four objects displayed on the monitor in response to spoken instructions presented in noise. They performed the speech recognition task while retaining for later recall either low (a single spoken digit) or high (four-digits) memory-load. In experimental trials, the named object shares phonology with the name of one of other presented objects. We compared eye-fixations on the named spoken target word, relative to fixations on its phonological competitor, as the word unfolded in time (online). Studies demonstrated that under adverse conditions, spoken word recognition dynamics differ significantly between situations in which the names of the target objects and competitors share an onset and those in which they share an offset in young adults

(McQueen and Huettig, 2012; Brouwer and Bradlow, 2016; Hadar et al., 2016), young adults with higher and lower working memory capacity (Nitsan et al., 2019), older adults (Ben-David et al., 2011), and hearing impaired listeners (McMurray et al., 2017). Therefore, the two types of phonological competition will be analyzed separately in the present study, and our analysis will focus on the onset overlap trials.

MATERIALS AND METHODS

Participants

Thirty-eight older adults were recruited from Reichman University's (IDC) older adult volunteer pool. Of this group, two were excluded due to loss of eye-tracking signal. Thus, the final group for analysis included 36 participants ($M_{\text{age}} = 67.9$ years, $SD = 3.2$, 20 females). All participants met the research inclusion criteria (see **Table 1** for details). Participants were paid 35 NIS (approximately \$10) for their participation. The number of participants was based on previous studies using a highly similar paradigm (Nitsan et al., 2019; Baharav et al., 2021).

Working Memory

Working memory span was assessed using the forward digit span subtest (Hebrew version of WAIS-III (*Goodman, 2001). To measure the participants' memory spans, sets of random digits were read aloud at a rate of one per second and they were instructed to repeat them, in the order in which they had been heard. The first list contained two digits, and the number of digits presented for recall increased gradually until the individual was no longer able to recall correctly. Two lists of each length were presented (e.g., two lists of three digits and then two lists of four digits, etc.). A single point was assigned to each list the participant correctly remembered (range of 0–16). Participants were divided into two subgroups based on their digit span scores (range 5–13). The lower-capacity subgroup consisted of 18 participants with a span score of five to nine ($M = 7.9$, $SD = 1.1$). The higher-capacity subgroup consisted of 18 participants with a span score of 10–13 ($M = 10.8$, $SD = 0.89$). The two groups did not differ in most individual characteristics, but differed on hearings status, with slightly better audiometric thresholds for the lower-capacity group (see **Table 2**).

Procedure

The experiment was administered individually in a dedicated sound attenuated booth (Iac Acoustics). Participants were seated 60 cm from a computer screen with their head placed in a customized chin rest to stabilize head movement. Each participant's dominant eye was calibrated to ensure that throughout the course of the trial participants' online eye-gaze position was recorded. A table mounted SR EyeLink 1000 eye-tracker in the “tower mount” configuration was used (SR Research Ltd., Kanata, ON, Canada). Eye-gaze position was recorded *via* the EyeLink software at a rate of 500 Hz.

During the experiment, two tasks were presented: spoken word recognition and digit recall (working memory load), conducted in a dual task situation. Trials began with a visual cue

TABLE 1 | Inclusion criteria for participant recruitment.

Inclusion criteria	
Language background	Proficient Hebrew speakers (no early bilinguals were included) assessed by a self-report and a score within the normal range in the WAIS-III Hebrew vocabulary subtest.
Hearing	Symmetrical air-conduction hearing thresholds expressed as pure tone averages (0.5, 1, and 2 kHz) of ≤ 25 dB HL in each ear, no reported history of auditory pathology. Audiometric assessment was conducted using a MAICO MA-51 audiometer using standard audiometric procedures in a sound attenuating testing booth.
Vision	Normal or corrected to normal visual acuity and color vision assessed by the Landolt-C charts and the Ishihara charts.
Cognition	Clinically normal scores for their age range on the MoCA cognitive screening test (≥ 22), and on the forward (≥ 5) and backward (≥ 4) digit span subtests (Hebrew version of WAIS-III; *Goodman, 2001).

TABLE 2 | Background information by working memory capacity group.

	Lower capacity	Higher capacity	Group comparison
<i>N</i>	18	18	
Age: mean (SD), years	68.5 (2.7)	67.5 (3.6)	$t = 0.95, p = 0.35$
Gender: count, females	9	11	$\chi^2 = 0.25, p = 0.62$
Hearing: mean (SD), 0.5, 1, and 2 kHz	15.1 (4.4)	18.7 (4.1)	$t = 2.5, p = 0.02$
Years of education: mean (SD)	16.5 (3.2)	16.2 (3.4)	$t = 0.25, p = 0.8$
MoCA: mean (SD)	25.5 (1.7)	26.3 (2.5)	$t = 1.2, p = 0.25$
Digit span: mean (SD)	7.9 (1.1)	10.8 (0.9)	$t = 8.5, p < 0.001$

of a black “play” triangle centered on the screen, immediately followed by the auditory presentation of the digit(s) preload through headphones, either one digit: low-load condition, or four digits: high-load condition. Participants were told to memorize these digits (in the order presented) for later recall. Then, a 3×3 grid with the four images would appear (**Figure 1A**). Participants were given 2 s to familiarize themselves with the four objects and their position on the computer screen. At the end of these 2 s a flickering fixation cross would appear in the center of the screen, once participants pressed the fixation cross to initiate the trial, the instruction sentence “point at the ___ [target word],” would be presented binaurally *via* the headphones. Selection of a named object was indicated by touching the object picture on the touch screen. Following the participant selection of a stimulus, a visual feedback signal: red highlight for an incorrect answer or green highlight for a correct answer, would appear in the square of the selected image. The visual display would then clear and a visual cue of a black circle would appear in the screen signaling participants to recall aloud the digit(s) preload from the beginning of the trial (**Figure 1B** illustrates the sequence of displays presented in each trial). The experimenter would then code the response (either correct or incorrect) online. Participants were instructed that speed and accuracy of both the object selection and digit recall were equally important. Participants completed 68 trials split into two trial blocks of the two memory load conditions (Low-load: one digit and High-load: four digits).

Each condition contained 34 trials of which two were practice trials, and 32 were experimental trials. The 32 trials in each condition were split such that 16 were “filler”: target object name did not share any phonology with the surrounding objects, and 16 were “critical” trials in which 8 were phonological onset competitors (e.g., /aɛ.nav/-/aɛ.gaz/ rabbit and box, respectively),

and 8 were phonological offset competitors (e.g., /xa.lon/-/ba.lon/ window and balloon, respectively).

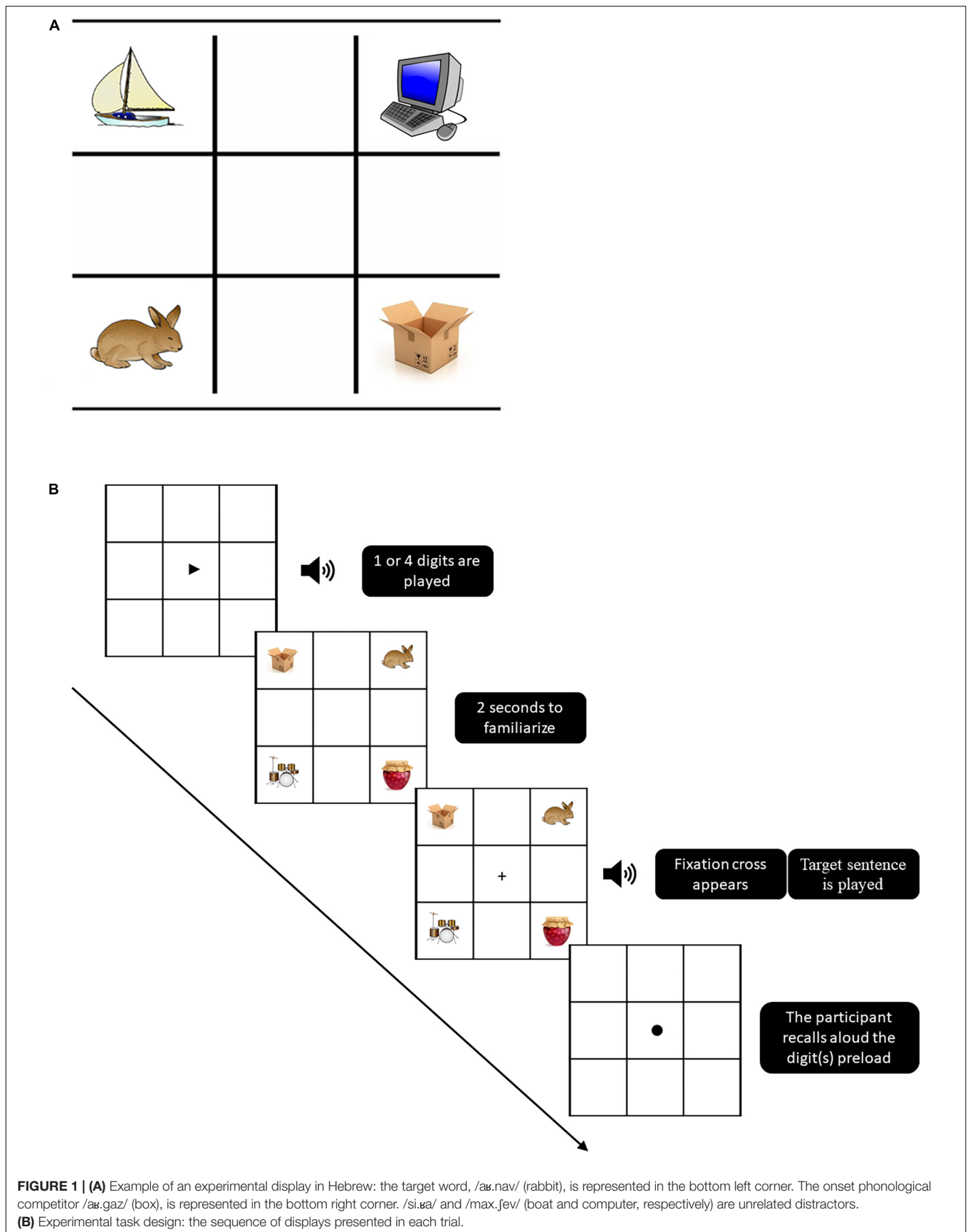
Stimuli

Auditory Stimuli

Stimuli were taken from Nitsan et al. (2019), and contained both the object names of the visual stimuli, and the sentence “point at the ___ [target word]” in Hebrew using a plural generic form. All object names were disyllabic. Average target word duration, including the Hebrew article *ha-* (the), was 1078 ms, SD = 91 ms (Nitsan et al., 2019). Considering that the definite article in Hebrew is not a separate word but a prefix, the target word onset was adjusted for each word separately (see Hadar et al., 2016). The root mean square (RMS) intensity was equated across all recorded sentences. Files were mixed with a continuous steady-state speech spectrum noise (for full details, see Ezzatian et al., 2010) at a fixed 0 dB SNR based off of values for discrimination timeline in Ben-David et al. (2012). Stimuli were presented binaurally at 50 dB above individual pure tone average (PTA) *via* a MAICO MA-51 audiometer using TDH 39 supra-aural headphones.

Visual Display

On each trial participants were presented with a 3×3 grid with four images of objects positioned at the grid corners. The stimuli (images) were previously used by Hadar et al. (2016), Nitsan et al. (2019), and Baharav et al. (2021) studies and were confirmed as clearly identifiable and highly familiar. In all trials one of the four image names represented the spoken target word and a second image’s name was a phonological competitor: sharing the initial syllable (onset sound overlap) or the final syllable (offset sound overlap) with the spoken target word. The remaining two objects presented on screen represented words that were phonologically and semantically unrelated to both the



target spoken word and phonological competitor. In critical trials the target word to be recognized was one of the two sound-sharing images. In addition to critical trials, filler trials were used to diminish participant expectation of phonetic resemblance between the words. Objects were presented twice during the experiment, once as a critical trial, and once as a filler trial in which one of the two phonologically “unrelated” items was used as the target word. To prevent implicit spatial learning, object positions on the screen were randomly rotated at each presentation (Farris-Trimble and McMurray, 2013).

Statistical Analysis

Growth curve analysis (GCA) (Mirman et al., 2008) was used to analyze the time course of fixation from word onset to 1200 ms after word onset (i.e., when target fixations had plateaued). To express listeners’ ability to discriminate the target word from its phonological competitor, we calculated *target discrimination scores* (following: Arnold et al., 2003; Kaiser and Trueswell, 2008; Brown-Schmidt, 2009; Ben-David et al., 2011). To generate the *target discrimination scores*, the proportion of fixations on the competitor was subtracted from the proportion of fixations on the target within 20 ms time bins, starting from the word onset to 1200 ms post word onset. In this measure, the higher the value the better listeners can discriminate the target from its phonological competitor; values approaching zero reflect an inability to discriminate between the target and competitor words. The overall time course of *target discrimination score* was captured with a second-order (quadratic) orthogonal polynomial with fixed effects of capacity group (low vs. high capacity) and working memory load (low vs. high load) on all time terms, and participant random effects on all time terms. The low working memory load condition and the high-capacity group was treated as the reference (baseline) and relative parameters estimated for the high working memory load condition and low-capacity group. These baseline conditions were selected to reflect preserved cognition and the easiest listening condition in this study. The two phonological competition conditions (onset and offset overlap) were modeled separately. Statistical significance (p -values) for individual parameter estimates was assessed using the normal approximation.

Offline response accuracy was analyzed using multilevel modeling (Heck et al., 2013) with fixed effects of capacity group (low vs. high capacity) and working memory load (low vs. high load) on response accuracy, participants were included as random effects. All analyses were carried out in SPSS version 25.

RESULTS

Onset Overlap – Accuracy of Behavioral Responses

Eye-gaze analysis included only trials in which participants both correctly selected the corresponding object on the visual display (indicating correct spoken word recognition) and correctly recalled the working memory load digits (indicating correct digit recall). **Table 3** shows mean accuracy performance across conditions and reflects differential effect of increased load for

TABLE 3 | Mean percentage (and SEs) of trials in which target word was correctly selected and digits were correctly recalled.

	Low WM capacity	High WM capacity
Low WM load	100% (0.0)	97.9% (1.13)
High WM load	83.8% (4.85)	95.1% (2.05)

Low and high working memory (WM) load, indicate the two preload conditions, one and four digit/s, respectively.

each working memory capacity group. In the low-capacity group, increasing memory load from one (low load) to four (high load) digits significantly reduced their response accuracy. However, the same increase in task demands did not change response accuracy for the high-capacity group. These differences were confirmed using a multilevel model as detailed in the statistical analysis section. The analysis revealed a main effect of load $F(1,34) = 13.21$, $p = 0.001$ on response accuracy and a significant interaction of load and span $F(1,34) = 6.60$, $p = 0.015$. LSD-corrected pairwise comparisons were conducted to clarify the interaction. It confirmed that the interaction of working memory load and capacity group was due to participants from the low-capacity group being significantly less accurate when a high load was present compared to when a low load present $F(1,34) = 19.25$, $p < 0.001$. In the high-capacity group accuracy did not differ significantly between the two load conditions $F(1,34) = 0.57$, $p = 0.456$.

Onset Overlap – Eye Gaze

The data and model fits are shown in **Figure 2**. Visual inspection of the left panel of **Figure 2A** shows that for listeners with lower working memory capacity, increasing task demands from low to high working memory load did not change the pattern and rate of target discrimination scores. In contrast, the right panel of **Figure 2B** indicates that for listeners with higher working memory capacity, increasing the working memory load delayed processing, suggesting less efficient spoken word processing. The results of the analysis as shown in **Table 4** confirm these observations. The analysis shows a significant effect of capacity group on the intercept and all polynomial time terms (linear and quadratic), suggesting that the rate of accumulating evidence from the unfolding spoken word differs between the two capacity groups. Working memory load was also found to have a significant effect on the linear and quadratic time terms, again suggesting a difference in evidence accumulation. Most importantly, the interaction between working memory load and capacity group on the linear and quadratic time terms was found to be significant.

A follow up model conducted separately for each capacity group revealed the source of this interaction (**Table 5**). In the low-capacity group, no significant effect of working memory load was evident; whereas in the high-capacity group the effect of working memory load on the linear and quadratic time terms was significant. The significant effect of working memory load on the linear term indicates a steeper slope, faster accumulation of evidence, under low working memory load. The effect of working memory load on the quadratic term further showcases

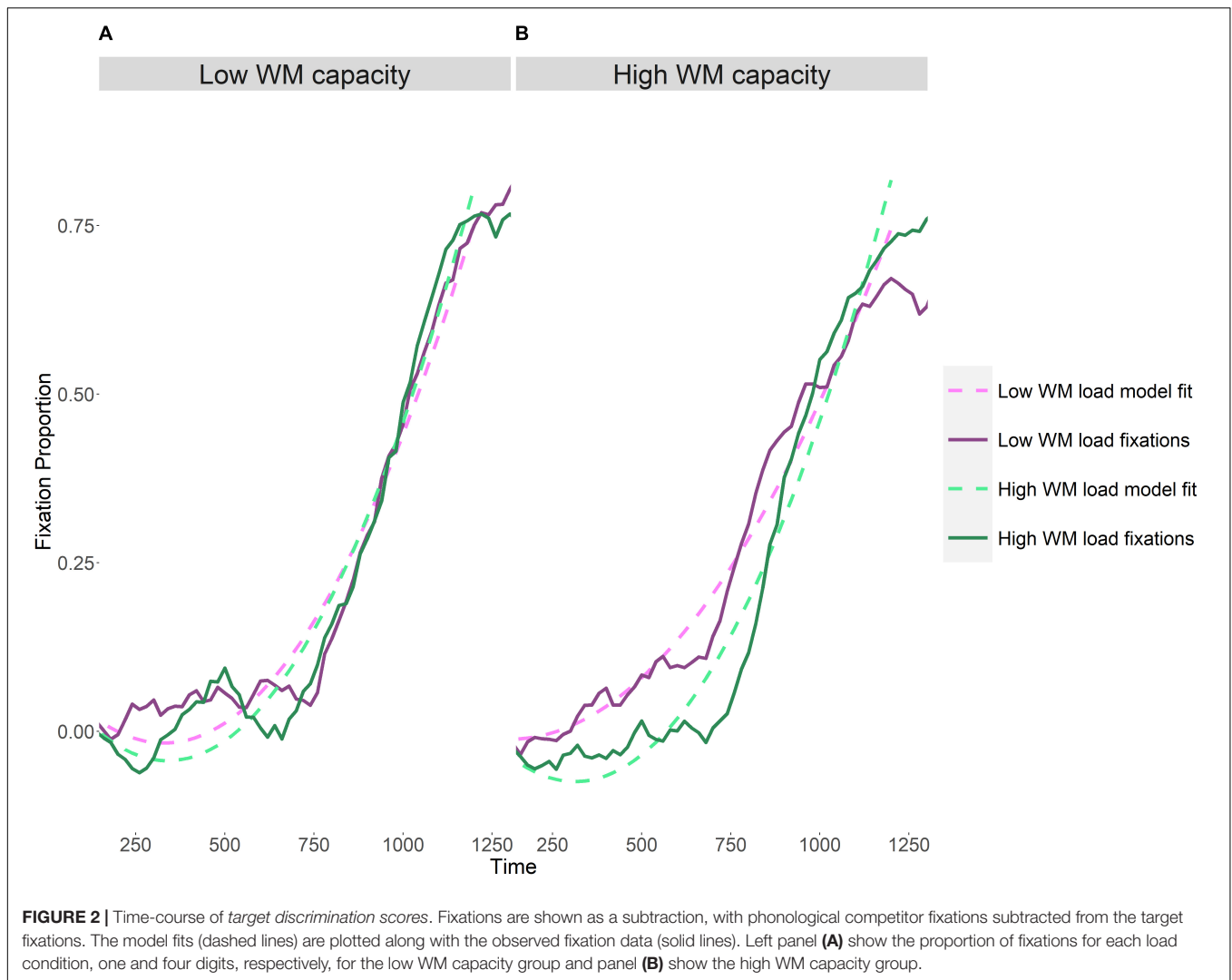


TABLE 4 | Results of growth curve analysis (GCA) – onset overlap.

	Term	Estimate	SE	t-Value	p
Participant group (WM capacity)	Intercept	0.094	0.041	2.28	0.025
	Linear	-0.502	0.105	-4.79	< 0.001
	Quadratic	0.004	<0.001	4.26	< 0.001
Working memory load	Intercept	0.037	0.027	1.37	0.170
	Linear	-0.541	0.105	-5.16	< 0.001
	Quadratic	0.001	<0.001	5.62	< 0.001
Participant group (WM capacity) × working memory load	Intercept	-0.041	0.038	-1.07	0.287
	Linear	0.429	0.148	2.89	0.004
	Quadratic	-0.001	<0.001	-2.86	0.004

a difference in the change in the rate of evidence accumulation between the two load conditions.

In sum, eye-movement analyses of onset overlap trials indicate that for the higher working memory capacity group, an increase in working memory load slowed spoken word processing. This slowdown was not evident for the lower working memory capacity group.

The same analyses conducted for the onset overlap trials were replicated for the offset overlap condition. The effects noted in the eye-gaze for the onset overlap condition were not found in the offset overlap, but for the effect of working memory capacity group. Analysis of accuracy of behavioral responses in the offset overlap revealed that increasing memory load from one to four digits significantly reduced listeners' response accuracy regardless

TABLE 5 | Results of growth curve analysis (GCA) conducted separately for each WM capacity group.

	Term	Estimate	SE	t-Value	p
Low-capacity group – working memory load	Intercept	−0.004	0.029	−0.12	0.902
	Linear	−0.112	0.112	−0.10	0.319
	Quadratic	0.000	0.000	1.47	0.141
High-capacity group – working memory load	Intercept	0.037	0.025	1.49	0.137
	Linear	−0.541	0.967	−5.59	<0.001
	Quadratic	0.000	<0.001	6.09	<0.001

of span group membership. Additionally, it shows that overall listeners from the high-capacity group had higher response accuracy compared to listeners from the low-capacity group. The low-capacity group had a greater reduction in response accuracy compared to the high-capacity group. The full analysis is provided in **Appendix A**.

DISCUSSION

We investigated the efficacy with which older adults with different working memory capacities process a spoken word in adverse conditions. Both online (eye-tracking) and offline (behavioral response accuracy) measures for spoken word recognition were used. Consistent with our hypothesis, we report that increasing task demands had different effects on listeners with higher vs. lower working memory capacity when the target and competitor shared onset sounds. Overall, listeners with higher working memory capacity were able to maintain their offline response accuracy at maximal performance even when they were asked to retain four digits for later recall instead of only one digit (high and low working memory load, respectively). However, this increase in working memory load had slowed down their online spoken word processing, suggesting less efficient processing at the single word level. For listeners with lower working memory capacity, increasing task demands significantly reduced offline recognition accuracy (from ~100 to ~80%), with no effect on online word processing. In the offset sound sharing condition, increasing memory load from one to four digits significantly reduced listeners' offline response accuracy regardless of their working memory capacity without affecting their online processing.

Our results present a clear support for the involvement of cognition, and more specifically working memory, in speech perception for older adults, even in the processing of a single spoken word. The literature to-date is inconsistent with regards to this question. Some studies on older adults observed correlations between working memory scores and recognition of single words in noise (Gordon-Salant and Cole, 2016; Heinrich and Knight, 2016) while others did not (Parbery-Clark et al., 2011; Smith and Pichora-Fuller, 2015). The present study has the distinct advantage of directly manipulating memory load, testing the effect of reduced cognitive resources on spoken word processing in aging. By varying the number of digits to be remembered (one vs. four digits) we were able to temporarily deplete spare cognitive capacity while listeners performed a speech recognition task in noise. This momentary depletion

led to changes in offline word recognition (for the lower-capacity group) and in online word processing (for the higher-capacity group). Note, if we were to test offline word recognition only, results would suggest that cognitive depletion mainly affects individuals with already low cognitive reserve. Indeed, previous works showed that increasing working memory load impairs language processing for clinical populations with reduced working memory capacity, such as people with aphasia, to a larger extent than for neurologically intact adults (Martin et al., 2012; Obermeyer et al., 2021). By using online measures, the current study shows the intricate effect of working memory depletion already at the single word level, even for individuals with larger cognitive reserves. Therefore, accessing and retrieving words from the mental lexicon when the input is degraded may require some available working memory resources even in healthy older adults with no signs of cognitive impairment. This link between cognition and speech processing in adverse listening conditions may stem from correlated activity across different brain regions. Indeed, spoken language processing rely on the joint activation of multiple cortical subsystems and several attempts were done to estimate its effectiveness by measuring cortical evoked responses (Gow, 2012). For example, Kim et al. (2021) suggested that changes in left supramarginal gyrus activity may be used as an independent predictor for speech processing efficiency.

In our analysis we found a differential effect of increasing working memory load for individuals with higher and lower working memory capacities. While increased load impaired offline accuracy for individuals with lower capacity, it affected online processing efficiency for individuals with higher capacity. According to the Framework for Understanding Effortful Listening (FUEL; Pichora-Fuller et al., 2016) speech processing depends on deployment of cognitive resources and therefore might be affected by differences in maximal capacity, especially under increased perceptual effort conditions such as in the presence of background noise and working memory load. It is possible that the listeners with lower working memory capacity were already using all their available resources in the low load condition in order to achieve maximal performance (100% accuracy). In other words, their online spoken word processing efficiency reflects their maximal ability. When facing increased task demands, they had no more available resources to allocate. Thus, with the same (maximal) word processing efficiency, as indicated by the online measures, their offline accuracy was significantly reduced. It is important to note that our analysis included only trials in which participants both correctly recognized the spoken word and correctly recalled the

working memory load digits. Removing incorrect trials arguably removes the most challenging trials from the analysis which might lead to an under-estimation of the effects of increased load on individuals with lower working memory capacity. In contrast, listeners with higher working memory capacity were not using all their available resources in the low load condition. Consequently, when working memory load increased they still had some spare available resources to allocate to maintain their performance. But this came with a cost of slower online word processing.

Our results might be interpreted in light of the Ease of Language Understanding (ELU) model (Rönnerberg et al., 2013). According to the model, understanding speech in adverse conditions is possible by drawing on central cognitive resources, mainly identified with working memory resources to compensate for the loss of automatic matching between the input and lexical representations when the input is degraded. Consistent with our findings, this model predicts that individuals with higher working memory capacity will be able to allocate these resources to maintain their offline performance. Changes in online processing could reflect either input degradation or the increased effort associated with the loss of automated word recognition.

In contrast to previous studies that relied on offline measures alone, the present study employed also online measures to track word processing as the acoustic signal unfolded over time. Standard measures of offline spoken word recognition accuracy do not capture the cost associated with maintaining a good level of performance. Our results highlight the importance of using both online and offline measures of speech processing to explore age-related changes in speech perception. The current study joins other studies that effectively used the visual world paradigm as a gauge of speech processing in adverse listening conditions (McQueen and Huettig, 2012; Helfer and Staub, 2014; Brouwer and Bradlow, 2016; McMurray et al., 2017). For example, McMurray et al. (2017) demonstrated that listeners with normal hearing process speech in a similar manner to that of cochlear implant users, when listening to severely degraded speech. In exploring the temporal dynamics of word recognition, authors could not only gauge the timing of target word recognition, but also determine the level and type of lexical competition that listeners were experiencing. Recent work from our lab also demonstrated that group-differences related to working memory load that were obscured in offline measures (e.g., accuracy) were uncovered when gaging online eye-tracking measures (Hadar et al., 2016; Nitsan et al., 2019; Harel-Arbeli et al., 2021).

Conclusions and Future Studies

The present data illustrate the differential effect of increasing task demands on spoken word recognition by listeners with higher vs. lower working memory capacity. Our findings suggest that additional cognitive capacity may lead to greater resilience of

older listeners to adverse listening conditions. Future studies may wish to examine this paradigm using different types of adverse listening condition such as fast speech. Understanding accelerated speech is another predominant complaint among elderly listeners but little is known about its time course (Humes and Dubno, 2010; Banai and Lavie, 2020; Rotman et al., 2020). Studies should also consider carefully controlling for the possible effects of stress and stereotype threat on hearing assessments (Ben-David et al., 2018; Nagar et al., 2022). Another path for investigation is testing these findings in clinical populations with cognitive decline (noting the difficulties in adaptation, Tziraki et al., 2017) or hearing aids and cochlear implant users to better tailor hearing rehabilitation expectations (e.g., Taitelbaum-Swead et al., 2022). Future studies may also choose to further examine the effects of working memory load and span on brain activity involved in speech processing in aging.

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 the School of Psychology Review Board, Reichman University, and by the Ethics Committee of the Faculty of Social Welfare and Health Sciences, University of Haifa. The participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

GN, KB, and BB-D wrote the manuscript. GN was responsible of the analysis and interpretation of the data. KB contributed to the conceptualizing of the research question and interpreting the results. BB-D was responsible of the design of the paradigm, the analysis, and the interpretation of the results. BB-D was the corresponding author and the study was conducted in his lab. All authors had a prominent intellectual contribution to the study, are accountable for the data and approved the final version of the manuscript.

ACKNOWLEDGMENTS

We wish to thank Dalith Tal-Shir, Tali Livne, and Shai Baharv for their help in recruiting participants and collecting data.

REFERENCES

- Abrams, L., and Farrell, M. T. (2011). "Language processing in normal aging," in *The Handbook of Psycholinguistic and Cognitive Processes: Perspectives in Communication Disorders*, eds J. Guendouzi, F. Loncke, and M. J. Williams (Hove: Psychology press), 49–73.
- Akeroyd, M. (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, 53–71. doi: 10.1080/14992020802301142
- Alloppenna, P. D., Magnuson, J. S., and Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: evidence for continuous mapping models. *J. Mem. Lang.* 38, 419–439. doi: 10.1006/jmla.1997.2558
- Arnold, J. E., Fagnano, M., and Tanenhaus, M. K. (2003). Disfluencies Signal Thee, Um, New Information. *J. Psycholinguist. Res.* 32, 25–36. doi: 10.1023/A:1021980931292
- Awh, E., and Vogel, E. K. (2008). The bouncer in the brain. *Nat. Neurosci.* 11, 5–6. doi: 10.1038/nn0108-5
- Ayasse, N. D., Lash, A., and Wingfield, A. (2017). Effort not speed characterizes comprehension of spoken sentences by older adults with mild hearing impairment. *Front. Aging Neurosci.* 8:329. doi: 10.3389/fnagi.2016.00329
- Baharav, S., Nitsan, G., and Ben-David, B. M. (2021). Commentary: working memory load affects processing time in spoken word recognition: test retest reliability of the E-WINDMIL eyetracking paradigm. *Front. Neurosci.* 15:663930 doi: 10.3389/fnins.2021.663930
- Baldwin, C. L., and Ash, I. K. (2011). Impact of sensory acuity on auditory working memory span in young and old adults. *Psychol. Aging* 26, 85–91. doi: 10.1037/a0020360
- Banai, K., and Lavie, L. (2020). Rapid perceptual learning and individual differences in speech perception: the good, the bad, and the sad. *Audit. Percept. Cogn.* 3, 201–211. doi: 10.1080/25742442.2021.1909400
- Ben-David, B. M., Chambers, C. G., Daneman, M., Pichora-Fuller, M. K., Reingold, E. M., and Schneider, B. A. (2011). Effects of aging and noise on real-time spoken word recognition: evidence from eye movements. *J. Speech Lang. Hear. Res.* 54, 243–262. doi: 10.1044/1092-4388(2010/09-0233)
- Ben-David, B. M., Erel, H., Goy, H., and Schneider, B. A. (2015). “Older is always better”: age-related differences in vocabulary scores across 16 years. *Psychol. Aging* 30, 856–862. doi: 10.1037/pag0000051
- Ben-David, B. M., Malkin, G., and Erel, H. (2018). “Ageism and neuropsychological tests,” in *Contemporary Perspectives on Ageism*, eds L. Ayalon and C. Tesch-Römer (Cham: Springer), 277–297. doi: 10.1007/978-3-319-73820-8_17
- Ben-David, B. M., Vania, Y. Y., and Schneider, B. A. (2012). Does it take older adults longer than younger adults to perceptually segregate a speech target from a background masker? *Hear. Res.* 290, 55–63.
- Besser, J., Koelewijn, T., Zekveld, A. A., Kramer, S. E., and Festen, J. M. (2013). How linguistic closure and verbal working memory relate to speech recognition in noise—A review. *Trends Amplif.* 17, 75–93. doi: 10.1177/1084713813495459
- Borovsky, A., Elman, J. L., and Fernald, A. (2012). Knowing a lot for one’s age: vocabulary skill and not age is associated with anticipatory incremental sentence interpretation in children and adults. *J. Exp. Child Psychol.* 112, 417–436. doi: 10.1016/j.jecp.2012.01.005
- Brouwer, S., and Bradlow, A. R. (2016). The temporal dynamics of spoken word recognition in adverse listening conditions. *J. Psycholinguist. Res.* 45, 1151–1160. doi: 10.1007/s10936-015-9396-9
- Brown-Schmidt, S. (2009). The role of executive function in perspective taking during online language comprehension. *Psychon. Bull. Rev.* 16, 893–900. doi: 10.3758/PBR.16.5.893
- Daneman, M., and Carpenter, P. A. (1980). Individual differences in working memory and reading. *J. Verb. Learn. Verb. Behav.* 19, 450–466.
- 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
- 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
- Ezzatian, P., Avivi, M., and Schneider, B. A. (2010). Do nonnative listeners benefit as much as native listeners from spatial cues that release speech from masking? *Speech Commun.* 52, 919–929. doi: 10.1016/j.specom.2010.04.001
- Farris-Trimble, A., and McMurray, B. (2013). Test–retest reliability of eye tracking in the visual world paradigm for the study of real-time spoken word recognition. *J. Speech Lang. Hear. Res.* 56, 1328–1345. doi: 10.1044/1092-4388(2012/12-0145)
- Fostick, L., Ben-Artzi, E., and Babkoff, H. (2013). Aging and speech perception: beyond hearing threshold and cognitive ability. *J. Basic Clin. Physiol. Pharmacol.* 24, 175–183. doi: 10.1515/jbcp-2013-0048
- *Goodman, L. (2001). Translation of WAIS-III - wechsler adult intelligence scale. *Psych. Tech.* 1, 133–136.
- Gordon-Salant, S., and Cole, S. S. (2016). Effects of age and working memory capacity on speech recognition performance in noise among listeners with normal hearing. *Ear Hear.* 37, 593–602. doi: 10.1097/AUD.0000000000000316
- Gow, D. W. (2012). The cortical organization of lexical knowledge: a dual lexicon model of spoken language processing. *Brain Lang.* 121, 273–288. doi: 10.1016/j.bandl.2012.03.005
- Hadar, B., Skrzypek, J. E., Wingfield, A., and Ben-David, B. M. (2016). Working memory load affects processing time in spoken word recognition: evidence from eye-movements. *Front. Neurosci.* 10:221 doi: 10.3389/fnins.2016.00221
- Harel-Arbeli, T., Wingfield, A., Palgi, Y., and Ben-David, B. M. (2021). Age-related differences in the online processing of spoken semantic context and the effect of semantic competition: evidence from eye gaze. *J. Speech Lang. Hear. Res.* 64, 315–327. doi: 10.1044/2020_JSLHR-20-00142
- Heck, R. H., Thomas, S. L., and Tabata, L. N. (2013). *Multilevel and Longitudinal Modeling with IBM SPSS*, 2nd Edn. Oxfordshire: Routledge, doi: 10.4324/9780203701249
- Heinrich, A., Henshaw, H., and Ferguson, M. A. (2015). The relationship of speech intelligibility with hearing sensitivity, cognition, and perceived hearing difficulties varies for different speech perception tests. *Front. Psychol.* 6:782. doi: 10.3389/fpsyg.2015.00782
- Heinrich, A., and Knight, S. (2016). “The contribution of auditory and cognitive factors to intelligibility of words and sentences in noise,” in *Physiology, Psychoacoustics and Cognition in Normal and Impaired Hearing*, eds P. van Dijk, D. Başkent, E. Gaudrain, E. de Kleine, A. Wagner and C. Lanting (Cham: Springer), 37–45. doi: 10.1007/978-3-319-25474-6_5
- Helfer, K., and Staub, A. (2014). Competing speech perception in older and younger adults: behavioral and eye-movement evidence. *Ear Hear.* 35, 161–170. doi: 10.1097/AUD.0b013e3182a830cf
- Humes, L. E. (2021). Factors underlying individual differences in speech-recognition threshold (SRT) in noise among older adults. *Front. Aging Neurosci.* 13:383. doi: 10.3389/fnagi.2021.702739
- Humes, L. E., and Dubno, J. R. (2010). “Factors affecting speech understanding in older adults,” in *The Aging Auditory System*, eds S. Gordon-Salant, R. D. Frisina, A. N. Popper, and R. R. Fay (Cham: Springer), 211–257. doi: 10.1007/978-1-4419-0993-0_8
- Humes, L. E., Watson, B. U., Christensen, L. A., Cokely, C. G., Halling, D. C., and Lee, L. (1994). Factors associated with individual differences in clinical measures of speech recognition among the elderly. *J. Speech Lang. Hear. Res.* 37, 465–474. doi: 10.1044/jshr.3702.465
- Janse, E. (2012). A non-auditory measure of interference predicts distraction by competing speech in older adults. *Aging Neuropsychol. Cogn.* 19, 741–758. doi: 10.1080/13825585.2011.652590
- Kaiser, E., and Trueswell, J. C. (2008). Interpreting pronouns and demonstratives in Finnish: evidence for a form-specific approach to reference resolution. *Lang. Cogn. Process.* 23, 709–748. doi: 10.1016/j.cognition.2009.03.010
- Kavé, G., and Halamish, V. (2015). Doubly blessed: older adults know more vocabulary and know better what they know. *Psychol. Aging* 30, 68–73. doi: 10.1037/a0038669
- Kim, S., Choi, I., Schwalje, A. T., Kim, K., and Lee, J. H. (2020). Auditory working memory explains variance in speech recognition in older listeners under adverse listening conditions. *Clin. Interv. Aging* 15, 395–406. doi: 10.2147/CIA.S241976
- Kim, S., Schwalje, A. T., Liu, A. S., Gander, P. E., McMurray, B., Griffiths, T. D., et al. (2021). Pre- and post-target cortical processes predict speech-in-noise performance. *Neuroimage* 228:117699. doi: 10.1016/j.neuroimage.2020.117699
- Lash, A., Rogers, C. S., Zoller, A., and Wingfield, A. (2013). Expectation and entropy in spoken word recognition: effects of age and hearing acuity. *Exp. Aging Res.* 39, 235–253. doi: 10.1080/0361073X.2013.779175
- Marslen-Wilson, W. (1990). “Activation, competition, and frequency in lexical access,” in *Cognitive Models of Speech Processing*, ed. G. T. M. Altmann (Cambridge, MA: The MIT Press), 148–172.

- Martin, N., Kohen, F., Kalinyak-Fliszar, M., Soveri, A., and Laine, M. (2012). Effects of working memory load on processing of sounds and meanings of words in aphasia. *Aphasiology* 26, 462–493. doi: 10.1080/02687038.2011.619516
- McMurray, B., Farris-Trimble, A., and Rigler, H. (2017). Waiting for lexical access: cochlear implants or severely degraded input lead listeners to process speech less incrementally. *Cognition* 169, 147–164. doi: 10.1016/j.cognition.2017.08.013
- McQueen, J. M., and Huettig, F. (2012). Changing only the probability that spoken words will be distorted changes how they are recognized. *J. Acoust. Soc. Am.* 131, 509–517. doi: 10.1121/1.3664087
- Meister, H. (2017). Speech audiometry, speech perception, and cognitive functions. *HNO* 65, 1–4. doi: 10.1007/s00106-016-0250-7
- Mirman, D., Dixon, J. A., and Magnuson, J. S. (2008). Statistical and computational models of the visual world paradigm: growth curves and individual differences. *J. Mem. Lang.* 59, 475–494. doi: 10.1016/j.jml.2007.11.006
- Nagar, S., Mikulincer, M., Nitsan, G., and Ben-David, B. M. (2022). Safe and sound: the effects of experimentally priming the sense of attachment security on pure-tone audiometric thresholds among young and older adults. *Psychol. Sci.* [Online ahead of print]. doi: 10.1177/09567976211042008
- Nitsan, G., Wingfield, A., Lavie, L., and Ben-David, B. M. (2019). Differences in working memory capacity affect online spoken word recognition: evidence from eye movements. *Trends Hear.* 23:233121651983962. doi: 10.1177/2331216519839624
- Norris, D. (1994). Shortlist: a connectionist model of continuous speech recognition. *Cognition* 52, 189–234. doi: 10.1037/0033-295X.115.2.357
- Obermeyer, J., Reinert, L., Kamen, R., Pritchard, D., Park, H., and Martin, N. (2021). Effect of working memory load and typicality on semantic processing in aphasia. *Am. J. Speech Lang. Pathol.* 31, 12–29. doi: 10.1044/2021_AJSLP-20-00283
- Parbery-Clark, A., Strait, D. L., Anderson, S., Hittner, E., and Kraus, N. (2011). Musical experience and the aging auditory system: implications for cognitive abilities and hearing speech in noise. *PLoS One* 6:e18082. doi: 10.1371/journal.pone.0018082
- Pichora-Fuller, M. K. (2008). Use of supportive context by younger and older adult listeners: balancing bottom-up and top-down information processing. *Int. J. Audiol.* 47, S72–S82. doi: 10.1080/14992020802307404
- Pichora-Fuller, M. K., Kramer, S. E., Eckert, M. A., Edwards, B., Hornsby, B. W., Humes, L. E., et al. (2016). Hearing impairment and cognitive energy: the Framework for Understanding Effortful Listening (FUEL). *Ear Hear.* 37, 5S–27S. doi: 10.1097/AUD.0000000000000312
- Pichora-Fuller, M. K., Schneider, B. A., and Daneman, M. (1995). How young and old adults listen to and remember speech in noise. *J. Acoust. Soc. Am.* 97, 593–608. doi: 10.1121/1.412282
- Rogers, C. S., and Peelle, J. E. (2021). “Interactions between audition and cognition in hearing loss and aging,” in *Speech Perception*, 1st Edn, eds L. Holt, J. Peelle, A. B. Coffin, A. N. Popper, and R. R. Fay (Cham: Springer International Publishing), doi: 10.31234/osf.io/d2bxw
- Rönnberg, 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
- Rönnberg, J., Rudner, M., Foo, C., and Lunner, T. (2008). Cognition counts: a working memory system for ease of language understanding (ELU). *Int. J. Audiol.* 47, S99–S105. doi: 10.1080/14992020802301167
- Rotman, T., Lavie, L., and Banai, K. (2020). Rapid perceptual learning: a potential source of individual differences in speech perception under adverse conditions? *Trends Hear.* 24:2331216520930541. doi: 10.1177/2331216520930541
- Smith, S. L., and Pichora-Fuller, M. K. (2015). Associations between speech understanding and auditory and visual tests of verbal working memory: effects of linguistic complexity, task, age, and hearing loss. *Front. Psychol.* 6:1394. doi: 10.3389/fpsyg.2015.01394
- Smith, S. L., Pichora-Fuller, M. K., and Alexander, G. (2016). Development of the word auditory recognition and recall measure: a working memory test for use in rehabilitative audiology. *Ear Hear.* 37, e360–e376. doi: 10.1097/AUD.0000000000000329
- Sommers, M. S. (2005). “Age-related changes in spoken word recognition,” in *The Handbook of Speech Perception*, eds D. B. Pisoni and R. Remez (Hoboken, NJ: Blackwell Publishing Ltd), 469–493. doi: 10.1002/9780470757024.ch19
- Sommers, M. S., and Danielson, S. M. (1999). Inhibitory processes and spoken word recognition in young and older adults: the interaction of lexical competition and semantic context. *Psychol. Aging* 14, 458–472. doi: 10.1037/10882-7974.14.3.458
- Stine-Morrow, E. A. L., Shake, M. C., Miles, J. R., and Noh, S. R. (2006). Adult age differences in the effects of goals on self-regulated sentence processing. *Psychol. Aging* 21, 790–803. doi: 10.1037/0882-7974.21.4.790
- Taitelbaum-Swead, R., Icht, M., and Ben-David, B. M. (2022). More than words: the relative roles of prosody and semantics in the perception of emotions in spoken language by postlingual cochlear implant users. *Ear Hear.* [Online ahead of print]. doi: 10.1097/AUD.0000000000001199
- Tziraki, C., Berenbaum, R., Gross, D., Abikhzer, J., and Ben-David, B. M. (2017). Designing serious computer games for people with moderate and advanced dementia: interdisciplinary theory-driven pilot study. *JMIR Serious Games* 5:e6514. doi: 10.2196/games.6514
- World Health Organization (2021). *Decade of Healthy Ageing: Baseline Report*. Available online at: <https://www.who.int/publications-detail-redirect/9789240017900> [accessed on January 14, 2021]
- Zekveld, A. A., Heslenfeld, D. J., Festen, J. M., and Schoonhoven, R. (2006). Top-down and bottom-up processes in speech comprehension. *Neuroimage* 32, 1826–1836. doi: 10.1016/j.neuroimage.2006.04.199

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.

Copyright © 2022 Nitsan, Banai and Ben-David BM. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). 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.

APPENDIX A. FULL ANALYSIS OF THE OFFSET OVERLAP CONDITION

Eye Gaze

Unlike the onset overlap condition, the analysis of offset overlap trials showed only an effect of span group on the linear and quadratic time terms, suggesting differential online word processing between span groups. **Table A1** summarizes the results of the analysis.

TABLE A1 | Results of growth curve analysis (GCA) – offset overlap.

	Term	Estimate	SE	t-Value	p
Participant group (WM capacity)	Intercept	-0.053	0.039	-1.37	0.174
	Linear	0.333	0.099	3.36	< 0.001
	Quadratic	<0.001	<0.001	-3.13	0.002
Working memory load	Intercept	-0.013	0.026	-0.51	0.607
	Linear	0.064	0.099	0.644	0.519
	Quadratic	<0.001	<0.001	-0.78	0.438
Participant group (WM capacity) × working memory load	Intercept	0.004	0.036	0.11	0.915
	Linear	-0.132	0.140	-0.94	0.347
	Quadratic	<0.001	<0.001	1.09	0.275

Accuracy of Behavioral Responses

Eye-gaze analysis included only trials in which participants both correctly selected the corresponding object on the visual display (indicating correct spoken word recognition) and correctly recalled the working memory load digits (indicating correct digit recall). The analysis indicated a main effect of load $F(1,34) = 34.23$, $p < 0.001$ and span group $F(1,34) = 6.83$, $p = 0.013$ on response accuracy. These two effects suggest that increasing memory load from one to four digits significantly reduced listeners' response accuracy regardless of span group membership. Additionally, it shows that overall listeners from the high span group had higher response accuracy ($M = 95.83$ vs. $M = 89.48$) compared to listeners from the low span group. The two effects interacted significantly $F(1,34) = 5.59$, $p = 0.024$. LSD-corrected pairwise comparisons revealed that increasing memory load yielded greater reduction in response accuracy in the low span group $F(1,34) = 33.75$, $p < 0.001$ compared to the high span group $F(1,34) = 6.08$, $p = 0.019$ as shown in **Table A2**.

TABLE A2 | Mean percentage (and SEs) of trials in which target word was correctly selected and digits were correctly recalled.

	Low WM capacity	High WM capacity
Low WM load	99.3% (0.69)	100% (0.0)
High WM load	79.7% (4.08)	91.7% (2.48)

Low and high WM, indicate the two preload conditions, one and four digit/s, respectively.