

Differences in Working Memory Capacity Affect Online Spoken Word Recognition: Evidence From Eye Movements

Trends in Hearing
Volume 23: 1–12
© The Author(s) 2019
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/2331216519839624
journals.sagepub.com/home/tia



Gal Nitsan^{1,2}, Arthur Wingfield³, Limor Lavie¹, and Boaz M Ben-David^{2,4,5}

Abstract

Individual differences in working memory capacity have been gaining recognition as playing an important role in speech comprehension, especially in noisy environments. Using the visual world eye-tracking paradigm, a recent study by Hadar and coworkers found that online spoken word recognition was slowed when listeners were required to retain in memory a list of four spoken digits (high load) compared with only one (low load). In the current study, we recognized that the influence of a digit preload might be greater for individuals who have a more limited memory span. We compared participants with higher and lower memory spans on the time course for spoken word recognition by testing eye-fixations on a named object, relative to fixations on an object whose name shared phonology with the named object. Results show that when a low load was imposed, differences in memory span had no effect on the time course of preferential fixations. However, with a high load, listeners with lower span were delayed by ~550 ms in discriminating target from sound-sharing competitors, relative to higher span listeners. This follows an assumption that the interference effect of a memory preload is not a fixed value, but rather, its effect is greater for individuals with a smaller memory span. Interestingly, span differences affected the timeline for spoken word recognition in noise, but not offline accuracy. This highlights the significance of using eye-tracking as a measure for online speech processing. Results further emphasize the importance of considering differences in cognitive capacity, even when testing normal hearing young adults.

Keywords

working memory, word recognition, online processing, eye-tracking, visual world paradigm

Date received: 21 March 2018; revised: 25 February 2019; accepted: 27 February 2019

Introduction

In the beginning of the study of individual differences in speech perception, research was mainly focused on differences in hearing thresholds as the main source for individual differences in peoples' ability to understand spoken language. However, over the past 25 years, it has become clear that although audiometric thresholds play a major role in predicting speech comprehension performance, it is not the only factor influencing this ability. The role of individual differences in various aspects of cognition has been gaining increased recognition as an important player in speech perception (in older adults: e.g., Dryden, Allen, Henshaw, & Heinrich, 2017; and in younger adults: e.g., Stenbäck, Hällgren, Lyxell, & Larsby, 2015). Within the realm of cognitive performance, working memory capacity is one of the most studied in predicting speech-in-noise performance.

Indeed, as speech processing was found to be costly in terms of working memory processing (e.g., Rönnberg, Rudner, Foo, & Lunner, 2008), it stands to reason that

¹Department of Communication Sciences and Disorders, University of Haifa, Israel

²Baruch Ivcher School of Psychology, Interdisciplinary Center Herzliya, Israel

³Volen National Center for Complex Systems, Brandeis University, Waltham, MA, USA

⁴Department of Speech-Language Pathology, University of Toronto, ON, Canada

⁵Toronto Rehabilitation Institute, University Health Networks, Toronto, ON, Canada

Corresponding author:

Boaz M. Ben-David, Communication, Aging and Neuropsychology Lab, Baruch Ivcher School of Psychology, Interdisciplinary Center Herzliya, P.O. Box 167, Herzliya 4610101, Israel.
Email: boaz.ben.david@idc.ac.il



differences in memory span will take a toll on spoken word processing itself. Füllgrabe and Rosen (2016) recently collected 24 data sets from the literature that tested this link with normal hearing young adults. In their meta-analysis, however, they note that individual differences in working memory capacity (as measured by the reading-span test scores) accounted for less than 2% of the variance in spoken-sentences in noise performance. This led the authors to conclude that for normal hearing young adults, differences in working memory capacity could not explain variation in spoken word identification accuracy. However, the debate on the role of individual differences in cognitive capacity in spoken word processing is not necessarily closed (e.g., see Akeroyd, 2008; Dryden et al., 2017).

Studies of the possible effects of individual differences in working memory on constraining the effectiveness of spoken word recognition have traditionally used offline measures such as giving an overt verbal response to presented words, after they have been heard, with measures such as the signal-to-noise-ratio (SNR) to reach 50% identification thresholds or performance accuracy at a fixed SNR level. By contrast, a recent study by Hadar, Skrzypek, Wingfield, and Ben-David (2016) took a different approach to this question by using an online, rather than an offline measure, to obviate the possibility that any working memory effect observed would not be the consequence of postword perception processing, or response initiation. Using the visual world eye-tracking paradigm (Cooper, 1974; Tanenhaus, Magnuson, Dahan, & Chambers, 2000; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995), they found that the timeline for spoken word recognition was slowed when listeners were required to retain in memory a list of spoken digits, while performing the recognition task. This hints that the link between individual differences in working memory capacity and speech perception may be evident during spoken word processing as the word unfolds in time. In the current study, we followed Hadar et al. (2016) using the visual world paradigm to examine the effect of differences in memory span on the time course of spoken word recognition in adverse listening conditions.

In its traditional use, the visual world eye-tracking paradigm asks listeners to follow spoken instructions referring to objects depicted on a computer monitor (the “visual world”), while the moment-to-moment location of their eye gaze is being recorded (Ben-David et al., 2011; Cooper, 1974; Tanenhaus et al., 1995, 2000). For example, participants might hear the sentence “touch the candy” while looking at a display containing four pictures, one representing the spoken target word (*candy*), one picture representing a phonological competitor (e.g., *cannon*), and two additional pictures of objects that are

neither semantically nor phonologically related to the target picture or its name (e.g., *table* and *zebra*). In Hadar et al.’s version of the paradigm, instructions were to touch a picture depicting the spoken word as quickly and accurately as possible, using a touch screen. By comparing the proportion of eye gazes on the target word over time with gazes on its competitors, this eye-tracking paradigm can reveal the time course for discriminating a target word from its phonological or other competitors.

It has been found in a number of studies that recording participants’ eye movements in relation to such a visual display provides a highly sensitive, continuous measure of the time course of spoken word processing (for a review, see Huettig, Rommers, & Meyer, 2011). This is so because the rapidity of one’s eye gaze on an object allows a nearly online determination of word recognition in contrast to estimates based on slower, offline measures, such as requesting a verbal or manual response (Ayasse, Lash, & Wingfield, 2017). That is, using this eye-tracking paradigm allows one to investigate the effect of different factors on the timeline for word recognition, not only its effects on recognition accuracy.

Hadar et al. (2016) used a Hebrew version of the visual world paradigm, in which the time course for word recognition was examined with a concurrent memory preload. Participants were asked to retain either one (lower load) or four (higher load) spoken digits for the duration of a word recognition trial. The rationale underlying the use of a digit retention preload rested on the assumption that holding a set of digits in memory draws on working memory resources that would otherwise be available for other processes (Baddeley & Hitch, 1974).

Although conceptions of working memory differ among theorists (cf. (Baddeley, 2012; Cowan, 1999; Engle, 2002; Oberauer, 2002), there is general agreement that the essence of working memory is represented by individuals’ limited capacity to retain information while manipulating this information in memory or while performing a concurrent task (cf. Baddeley & Hitch, 1974; McCabe, Roediger, III., McDaniel, Balota, & Hambrick, 2010; Postle, 2006). As a consequence, in Hadar et al.’s (2016) case, an increase in a demand on working memory capacity by requiring retention of four spoken digits rather than one would be assumed to decrease the capacity available to perform other tasks, possibly including spoken word recognition. That is, considering that successful word recognition requires working memory resources, one would expect a larger memory preload to interfere with the timeline for recognition more than a smaller memory preload would (Baddeley & Hitch, 1974). Indeed, Hadar et al. (2016) found that the time of eye fixations on a named object

relative to nontarget objects was slowed by the larger memory preload.

These findings were taken to suggest that cognitive load affects the online processing of spoken words, even when performed by normal-hearing, cognitively able young adults. To the extent that their conclusion is correct, we would predict that working memory preload, as manipulated by Hadar et al., would be affected by an individual's memory span. Specifically, one might expect that the difference between individuals with lower and higher memory spans will be revealed when memory preload is high.

The Current Study

Hadar et al.'s study could not present evidence on the role of individual differences in memory span on spoken word recognition. They presented spoken words in quiet, which resulted in a ceiling effect (100%) in accuracy, where individual differences might be obscured as the task may not exceed maximal capacity. In the current study, we expanded on their conclusions by testing differences in memory span. Using the visual world paradigm, we compared participants with higher and lower memory spans on the time course for spoken word recognition when available working memory resources were reduced by presentation of a memory preload. To avoid potential ceiling effects, we presented words in noise at a fixed level of SNR to yield ~80% accuracy. As a result, we could test the effects of differences in memory span on the timeline for spoken word recognition separately from accuracy. Whether or not differences in memory span further affect the timeline for word recognition in these conditions would offer a strong test for the influence of individual cognitive abilities on online processing of spoken word recognition.

Method

Participants

Thirty-seven young adults were recruited from the Interdisciplinary Center (IDC) Herzliya in return for partial course credits. Of this group, six were excluded: Three participants failed to follow the instructions, and the data for another three participants were lost due to a failure in eye-movement recording. Thus, the final group for analysis included 31 participants (M age = 24.6 years, $SD = 3.3$). All participants had normal or corrected-to-normal vision, as tested using Landolt-C charts for near vision, and when necessary used their own corrective eyewear.

Audiometric assessment was conducted using a MAICO MA-51 audiometer using standard audiometric procedures in a sound attenuating testing booth.

All participants had pure-tone air conduction thresholds within clinically normal limits measured across 250 to 8000 Hz in both ears (≤ 15 dB HL). The "groups" mean pure tone average across .5, 1, and 2 kHz was 7.1 dB HL ($SD = 4.4$).

All participants were native Hebrew speakers, as assessed by a detailed questionnaire and a score corresponding to an above average level for native Hebrew speakers ($M = 47.4$, $SD = 5.6$) on the Wechsler vocabulary subtest (Hebrew version of WAIS-III, Goodman, 2001; for a discussion on WAIS vocabulary subset and language proficiency, see Ben-David, Erel, Goy, & Schneider, 2015). In addition, all participants were interviewed by the first author, a registered speech-language pathologist, to confirm their proficiency in Hebrew as well as their language background (native speakers, who spoke only Hebrew at home or work). No early bilinguals were included in the study.

Participants' memory spans were assessed by presenting aloud sets of random digits at a rate of one per second, with instructions to report them back verbatim, in the order in which they were heard. The shortest list contained two digits, with the number of digits presented for recall increasing progressively until the individual was no longer able to recall all of the digits accurately and in the correct order. Participants received two lists of each length (e.g., two lists containing three digits and then two lists of four digits, etc.), with the individual's span taken as the maximum list length at which at least one of the two lists was accurately recalled (Lezak, 1995).

Participants were divided into two subgroups based on their digit span scores. Individuals assigned to the lower span subgroup consisted of 15 participants that preformed in the lower half of the range of the digit span scores, with a span of five to six digits ($M = 5.6$, $SD = .45$). Sixteen participants assigned to the higher span subgroup preformed in the upper half of the range, with a span of seven to eight digits ($M = 7.3$, $SD = .46$). The two groups did not differ significantly in hearing acuity, vocabulary, age, $t(29) < 1.60$, $p > .05$ (for all tests), or gender, $\chi^2(1, N = 31) = 0.35$, $p = .55$. This grouping follows the method suggested by others (e.g., Gordon-Salant & Cole, 2016), acknowledging that a difference of one digit may not be revealing (as the variances in performance scores for this age range, documented by the Hebrew version of WAIS-III, Goodman, 2001, were 1.27–1.35 digits).

Stimuli

Auditory stimuli. The auditory stimuli were taken from Hadar et al. (2016) and consisted the Hebrew equivalent of the sentence "point at the ____ [target word]" using the plural nongender specific form (i.e., "/hats.bi.u/ /al/ /

ha[target word]). All target words were common, disyllabic names of picturable objects. The instruction (“point at the”) and the object name were digitally recorded by a native Hebrew speaking radio actress in a professional radio studio (IDC radio) using a sampling rate of 48 kHz. The root-mean-square intensity was equated across all recorded sentences. The average target word duration (including the Hebrew definite article *ha-*) was 1078 ms ($SD = 91$ ms).

The speech stimuli (the complete sentences of the form “point at the ___” in Hebrew) were mixed with continuous steady-state speech spectrum noise (for full details, see Ezzatian, Avivi, & Schneider, 2010) at a fixed -4 dB SNR, following Ben-David et al., 2011; Ben-David, Tse, & Schneider, 2012. This SNR level was found to elicit approximately 80% accuracy in a pretest conducted using these stimuli. Materials were presented binaurally at 79 dB SPL via a MAICO MA-51 audiometer using TDH 39 supra-aural headphones.

Visual displays. For each trial, participants saw four pictures of objects displayed in the four corners of a 3×3 grid on a touch screen computer monitor (T 23" ATCO infrared 4096×4096). Object pictures were taken from the normed color image set of Rossion and Pourtois (2004), supplemented by images from commercial clip art data bases selected to match the Rossion and Pourtois images in visual style. All stimuli were taken from Hadar et al.'s study (2016). A pretest confirmed that all visual displays were clearly identifiable and that all depicted nouns were highly familiar (for details, see Hadar et al., 2016).

In all cases, the name of one of the depicted objects matched the target word. In the critical trials, a second object on the computer screen was a phonological competitor: An object whose name shared either the initial syllable (onset sound overlap) or in the final syllable (offset sound overlap), with the remaining two objects unrelated either phonologically or semantically to either the target word or the phonological competitor.

An example display illustrating a critical trial with an onset phonology overlap is shown in Figure 1 for the target word /aʁ.nav/ (rabbit). The target picture of a rabbit is shown in the lower left corner, an onset phonological competitor /aʁ.gaz/ (box) is shown in the bottom right corner, and two phonologically and semantically unrelated distractors, /si.ba/ and /max.fev/ (boat and computer, respectively), are shown in the upper right and left corners.

Filler trials were also employed in order to diminish participants' expectations about the task and a phonetic resemblance between the target and other objects in the display. In these trials, none of the distractors shared either phonological onsets or phonological offsets with

the target word. None of the object pictures were reused in either the critical or filler trials.

Procedure

Each participant received a total of 64 trials, 16 with a phonological onset competitor, 16 with a phonological offset competitor, and 32 filler trials in which none of the competitor object names shared any phonology with the target word. Instructions to the participants were to listen to the instruction sentence and target word (e.g., “Point at the *rabbit*”), and to use their index finger to point at the named object by touching the object picture on the computer screen. Both accuracy and speed of the pointing response were requested.

Participants were told that prior to each trial, they would hear either one digit (low-load condition) or a set of four digits presented at a rate of one digit per second (high-load condition). They were told that their task would be to remember the digit(s) and to recall the one digit or four digits aloud, when prompted, after they had heard the carrier sentence and target word, and had made their pointing response.

Participants were seated 60 cm from the computer screen with their head placed in a chin rest to stabilize head movement. Throughout the course of each trial, the participant's moment-to-moment eye gaze position on the computer screen was recorded via a table-mounted SR Eyelink 1000 eye-tracking apparatus using the “tower mount” configuration (SR Research Ltd., Ontario, Canada). Eye gaze position was sampled at a rate of 500 Hz and recorded via Eyelink software.

Each trial began with a visual alerting cue (a black triangle centered on the computer screen), immediately followed by the auditory presentation of the digit(s) preload. The preload presentation was then followed by the 3×3 grid appearing on the computer screen containing four pictured objects at each corner of the grid. Participants were allowed 2 s to view the objects and their positions on the screen. Following this familiarization period, participants received a brief 1000 Hz tone that signaled the participant to focus on a fixation cross that simultaneously appeared in the center of the grid.

After the system registered cumulative fixations on the central square for at least 200 ms, the fixation cross disappeared, and the recorded instruction sentence was presented. A feedback signal followed the participant's selection. This took the form of a green square denoting a correct response or a red square denoting an incorrect response appearing over the participant's selected object. The visual display then cleared, and a visual cue (a black circle on a white background) appeared signaling the participant to recall the digit preload. Digit(s) recall was given aloud and coded online by the experimenter.

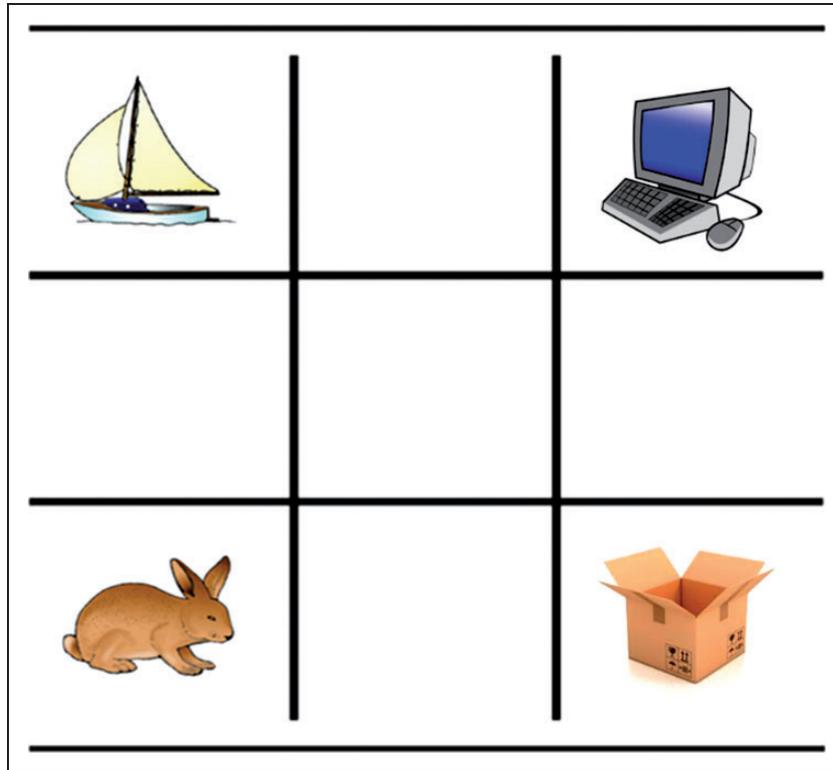


Figure 1. Example of the experimental display. The target word, in this example, /æ.nav/ (rabbit), is represented in the bottom left corner. The phonological competitor /æ.gaz/ (box) is represented in the bottom right corner. /si.ɒl/ and /max.jev/ (boat and computer, respectively) are unrelated distractors.

High- and low- preload conditions were separately blocked, with the order of preload blocks (high and low) counterbalanced between participants. The onset overlap, offset overlap, and filler conditions were intermixed in presentation within each preload block. A sole restriction was that the first four trials in each block were always fillers. The relative positions of the target, the phonological competitor, and the unrelated pictures within the grid displays were counterbalanced across the set of displays. To control for word frequency effects, target–competitor words allocation was counterbalanced, such that each word served for half of the participants as a target and for the other half as a phonological competitor and vice versa. The main experiment was preceded by a practice session with eight preload and sentence sets to familiarize participants with the task and instructions.

Results

Response Accuracy

Table 1 shows, for each of the experimental conditions, the percentage of trials in which participants both correctly selected the corresponding object on the visual

Table 1. Mean Percentage (and Standard Deviations) of Trials in Which Target Word Was Correctly Selected and Digits Were Correctly Recalled.

	High WM load	Low WM load
Onset-sound sharing		
Higher span	82.5% (16.2)	83.3% (12.2)
Lower span	79 % (13.5)	82.3% (15.5)
Offset-sound sharing		
Higher span	85.8% (11.4)	85.9% (10.4)
Lower span	83.3% (11.1)	84.4% (17.7)

WM = working memory.

display (indicating correct spoken word recognition) and correctly recalled the preload digits. The similarity in accuracy across the experimental conditions as seen in Table 1 was reflected in a 2 (Working Memory Load: high, low) × 2 (Competitor Type: onset overlap, offset overlap) × 2 (Participant Group: higher span, lower span) mixed design analysis of variance that failed to reveal significant main effects of working memory load, competitor type, or memory span group, nor any significant interactions ($p > .05$).

Eye Gaze Analysis

Growth curve analysis (Mirman, 2014; Mirman, Dixon, & Magnuson, 2008; using an R statistics packages, see details in Appendix A) was used to analyze the target gaze data from word onset to 3500 ms after word onset. Only trials in which participants both correctly selected the corresponding object on the visual display (indicating correct spoken word recognition) and correctly recalled the preload digits were included in the analysis (removing an average of 16.7% of the trials, across groups and conditions). Saccades and fixations on grid cells other than the four pictograms and central fixation cross, as well as eye blinks, were coded as events outside the interest areas. The eye gaze data were aggregated in nonoverlapping 20 ms time bins, with 10 samples summed per time bin (cf. Arnold, Fagnano, & Tanenhaus, 2003; Ben-David et al., 2011; Brown-Schmidt, 2009; Kaiser & Trueswell, 2008). The overall time course of target fixations was modeled with a third-order (cubic) orthogonal polynomial and fixed effects of: Working memory load (low vs. high; within-participants), Competitor type (onset vs. offset overlap; within-participants), and Participant group (higher vs. lower span; between participants) on all time terms and the intercept, as well as fixed effects corresponding to the

3 two-way and 1 three-way interactions between these variables. The model also included participant random effects on all time terms and participant-by-condition random effects on all time terms except the cubic (for details see Mirman et al., 2008). The model fits (lines) are shown in Figure 2 along with the mean observed target fixation data (symbols). We note that proportion of fixations on the unrelated items were minuscule (lower than 5% across the entire data set).

The proportion of fixations on the phonological competitors was also minor (not exceeding 10%). Attempts to fit the model to this data set yielded insufficient fit parameters. A 2 (Working Memory Load: high, low) \times 2 (Competitor Type: onset overlap, offset overlap) \times 2 (Participant Group: higher span, lower span) mixed design analysis of variance revealed no significant main effects of working memory load, competitor type, or memory span group, nor any significant interactions on competitors fixations ($p > .05$). For a graphic description, see online Appendix B.

Table 2 presents the results of the model, and Table 3 presents the mean thresholds in ms for 50% and 75% as derived from the model. There was a significant effect of working memory load on the intercept term, indicating lower overall target fixation proportions for the high-load condition relative to the low-load

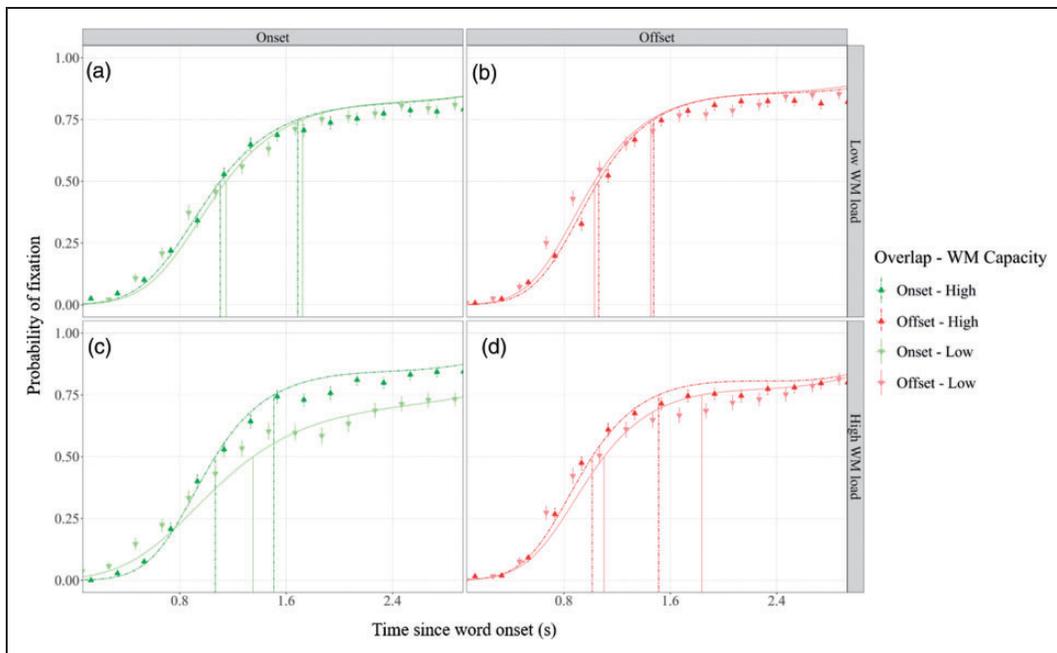


Figure 2. Mean proportion of fixations to the target (with SE bars) for participants with higher and lower memory spans. Top panels show the fixations when there was a low (one digit) working memory load, in onset competition (Panel A) and offset competition (Panel C). The bottom panels show the data for a high (four digit) working memory load, in onset competition (Panel B) and offset competition (Panel D). The model fits (lines) are plotted along with the observed target fixation data (symbols). The vertical lines represent the 50% and 75% thresholds (dashed and solid lines, respectively). WM = working memory.

Table 2. Results of Growth Curve Analysis—Target Fixations Model.

Term	Estimate	Standard error	Z	p <
Working memory load				
Intercept	−0.040	0.004	−9.65	.001
Linear	−3.133	0.073	−42.57	.001
Quadratic	2.056	0.067	30.36	.001
Cubic	−0.754	0.057	−13.08	.001
Competitor type				
Intercept	0.301	0.019	15.52	.001
Linear	0.547	0.355	1.54	ns
Quadratic	−0.122	0.322	−0.38	ns
Cubic	0.853	0.270	3.15	.01
Participant group (span)				
Intercept	0.061	0.304	0.20	ns
Linear	−6.188	4.625	−1.34	ns
Quadratic	3.692	2.327	1.59	ns
Cubic	−1.106	1.637	−0.68	ns
Participant Group (Span) × Working Memory Load				
Intercept	0.033	0.005	5.72	.001
Linear	4.890	0.103	47.03	.001
Quadratic	−3.128	0.095	−32.76	.001
Cubic	1.410	0.080	17.54	.001
Competitor Type × Working Memory Load				
Intercept	−0.057	0.006	−8.88	.001
Linear	2.166	0.117	18.47	.001
Quadratic	1.537	0.106	−14.42	.001
Cubic	0.946	0.090	10.50	.001
Competitor Type × Participant Group (Span)				
Intercept	−0.319	0.025	−12.57	.001
Linear	8.024	0.460	17.44	.001
Quadratic	−5.367	0.419	−12.79	.001
Cubic	1.099	0.350	3.13	.01
Competitor Type × Working Memory Load × Participant Group (Span)				
Intercept	0.068	0.008	7.94	.001
Linear	−5.374	0.156	−34.39	.001
Quadratic	3.167	0.142	22.17	.001
Cubic	−1.315	0.119	−10.97	.001

ns = not significant.

($Estimate = -.04$, $SE = .004$, $p < .001$), that is, a delay in target-fixations in high-load trials. A significant effect of working memory load on the linear term was found, indicating slower accumulation of evidence rate for the high-load condition. Significant effects of working memory load were also indicated on the quadratic term and cubic terms, further showcasing that the increase in working memory load slowed spoken word recognition processes (shallow curvature; see Table 2 for full results).

Table 3. Thresholds (in Milliseconds) Derived From the Growth Curve Analysis Model for 50% and 75% Target Fixations, as a Function of the Type of Phonological (Onset vs. Offset) Overlap, Working Memory Load (High vs. Low), and Memory Span Group (Higher vs. Lower).

	High WM load		Low WM load	
	50%	75%	50%	75%
Onset-sound sharing				
Higher span	1,070	1,510	1,100	1,690
Lower span	1,350	3,010	1,150	1,720
Offset-sound sharing				
Higher span	1,010	1,510	1,060	1,470
Lower span	1,100	1,840	1,030	1,450

WM = working memory.

A significant effect of competitor type was evident on the intercept term, indicating higher overall target fixation proportions for the offset overlap relative to the onset overlap ($Estimate = .30$, $SE = .02$, $p < .001$). That is, reflecting a delay in target fixations in onset overlap trials, across participant groups and working memory load conditions. There was also a significant effect of competitor type on the cubic term, but all other effects of competitor type were not significant (see Table 2 for full results). The effect of competitor type interacted significantly with working memory load and it also interacted significantly with participant group, on all terms (Lines 5 and 6 in Table 2). This suggests that the delay in target fixation proportions for the onset overlap trials was moderated by participant span group and working memory load conditions.

No significant main effects were found for the participant group (higher vs. lower span), on of the terms (intercept, linear, quadratic, or cubic). Importantly, significant interactions were found for participant group and working memory load on all terms (see Line 4 in Table 2). Examining Figure 2, these interactions suggest that a delay in target fixations and slowing in accumulation rates for the higher span relative to the lower span group was only evident when a high working memory load was imposed.

A significant three-way interaction of competitor type, working memory load, and participant group was evident on all terms (Line 7 in Table 2). Examining Figure 2, it appears that the three-way interaction reflects that the extent of the effects of participant span group on online speech processing when a high load was imposed (i.e., delay, reduced rate of data accumulations, etc.) was larger for onset overlap trials, as compared with offset overlap trials. This is also indicated in the effect of participant group on 50% thresholds for high-load trials, as presented in Table 3. For example, in the high-load

condition, we note a delay of 90 ms comparing individual with lower- and higher span, in offset overlap trials. This span-related delay inflated to 280 ms in onset overlap trials.

General Discussion

In the current study, we tested the effects of differences in memory span on the timeline for spoken word recognition. We compared two groups of normal-hearing young adults who differed in memory span, as indexed by their forward digit spans. Using the eye-tracking visual world paradigm, listeners were asked to follow spoken instructions while retaining either a low (single-digit) or high (four-digit) load for later recall. In critical trials, instructions (“point at the ___”) directed listeners’ gaze to a named object on a visual display that shared either onset or offset sounds with a displayed competitor. Our results show no differences in performance accuracy across conditions and participant groups (see Table 1). These accuracy findings are consistent with Füllgrabe and Rosen’s (2016) review of the literature that suggested that individual differences in working memory capacity could not explain variation in spoken word identification accuracy for normal-hearing young adults. However, in our analysis of eye-movements, significant effects for span differences were evident in the timeline for spoken word processing.

An advantage of eye gaze as an index of the time course of spoken word recognition lies in its rapidity and its independence from the potential confound of motor speed that can affect traditional overt verbal or keypress recognition responses (Ayasse et al., 2017). Following this principle, Hadar et al. (2016) used high- and low-preload conditions, analogous to the present experiment, to demonstrate that a higher working memory load slowed participants’ relative fixation time on an object representing the spoken target word. In the current study, we introduced the variable of memory span by comparing participants who were tested to have higher versus lower span scores. This follows an assumption that the interference effect of a memory preload is not necessarily a fixed value. Rather, its effect will be greater for individuals with a smaller memory span, with holding span representing a critical component of the full working memory system (Baddeley, 2012). As we saw, this expectation was confirmed by a finding that participants’ memory spans had little to no effect on the time course of preferential eye fixations on the named object when there was a low working memory load. By contrast, when the task included a high working memory load, individuals’ memory spans had a significant effect on the time course of word recognition as measured by preferential eye gaze. Indeed, when a high preload was imposed, the timeline for fixations on the

target word was delayed by ~550 ms for listeners with lower memory span, relative to higher span listeners (see Table 3), and the rate of evidence accumulation was affected. This delay was not evident in the low preload condition.

This discrepancy between the impact of participants’ memory span on spoken word online processing, and the absence of an effect on recognition accuracy, may appear paradoxical. A possible explanation is that span differences might be absorbed by the time it takes participants to initiate the offline response (~2 s postword onset). This rationale echoes Ben-David et al.’s (2011) data showing that when group effects were apparent comparing older and younger adults, they were manifested by a delay in the timeline for target fixations, but not by a difference in accuracy.

As noted earlier, the role of individual differences in working memory on speech perception is debated in the literature (cf. Akeroyd, 2008; Dryden et al., 2017; Füllgrabe & Rosen, 2016). Although a number of offline tests of speech recognition generally suggest that individual differences in memory span may not have a large effect on performance for young adults (Füllgrabe & Rosen, 2016), other offline studies do (Gordon-Salant & Cole, 2016). An example of the latter can be seen in Benichov, Cox, Tun, and Wingfield (2012), who found that individual differences in a cognitive composite based on episodic memory, working memory, and speed of processing contributed to ease of spoken word recognition (as indexed by the minimum SNR that allowed for correct word recognition). This cognitive factor had an effect, even when spoken words were presented without a linguistic context, and with individual differences in hearing acuity and verbal ability taken into account.

Spoken language processing involves holding and integrating phrases and clauses to create a coherent representation of its meaning. This is assumed to be largely supported by working memory resources (Wingfield & Tun, 2007), engendering the impact of individual differences in working memory on predictive language processes. Indeed, evidence from an online eye-tracking study by Huettig and Janse (2016) suggest that individual differences in working memory (a combined score of digit span, spatial and auditory tests) may affect anticipatory language processing that occurs before the target word is heard (e.g., processing linguistic, semantic, and environmental context). We suggest here that differences in working memory span can also affect spoken word processing itself, as the word unfolds in time. Specifically, even a relatively minor load of memorizing four spoken digits was sufficient to reveal large differences in the timeline for processing spoken words in adverse listening conditions (SNR = -4 dB) between individuals with lower and higher memory spans. It is notable that the effects of span differences emerged only

when a high load was imposed. This offers strong evidence for the influence of cognitive abilities on the online processing of spoken word recognition in adverse conditions.

Our conclusion joins others showing effects of working memory on speech recognition performance in a range of listening conditions (e.g., Besser, Koelewijn, Zekveld, Kramer, & Festen, 2013; Daneman & Merikle, 1996; Lash & Wingfield, 2014; Rudner, Lunner, Behrens, Thorén, & Rönnberg, 2012; Sörqvist & Rönnberg, 2012). In the current study, we took advantage of the short latencies of eye-movements to show that a combination of imposed (extrinsic) working memory demands and individual (intrinsic) differences in memory spans combine in their effects on the time course of word recognition even in the absence of a constraining linguistic context. Thus, the current study can be taken as a further support to the evidence showing that spoken word recognition in adverse conditions is not a resource-free process (see Akeroyd, 2008; Dryden et al., 2017), but rather, taps into cognitive resources, at least to maintain the activation of lexical candidates (Zhang & Samuel, 2018).

In the present study, we observed detrimental effects on the time course of word recognition from both phonological onset and phonological offset competition, even if to a larger extent in onset overlap trials. It is assumed that as a spoken word is presented, initial competition will be created by words sharing the same onset sounds as the target word, but that competition will also arise from words sharing offset phonology as the target word unfolds in time (cf. Luce & Pisoni, 1998; Marslen-Wilson, 1990; Sommers & Amano, 1998; Wayland, Wingfield, & Goodglass, 1989; Wingfield, Goodglass, & Lindfield, 1997).

Past studies using the visual world paradigm to explore relative interference effects from words sharing onset versus offset phonology in ideal listening conditions have yielded mixed findings. Whereas some studies found effects mostly for onset overlap (e.g., the quiet condition in McQueen & Huettig, 2012), some found effects mostly for offset overlap (e.g., Hadar et al., 2016), and others found effects for both types of overlap (Alloppenna, Magnuson, & Tanenhaus, 1998). However, when this was tested in adverse listening conditions (e.g., words in background noise), the extent of lexical competition from offset sound sharing alternatives has increased (Ben-David et al., 2011; Brouwer & Bradlow, 2016; McQueen & Huettig, 2012). For example, when McQueen and Huettig (2012) replaced some phonemes in a carrier sentence by noise, the proportion of fixations on the offset overlap items increased. Similarly, when Brouwer and Bradlow (2016) simply presented spoken words on the background of broadband noise, the ratio of fixations on the offset overlap competitor increased, as compared with the quiet condition.

Brouwer and Bradlow (2016) suggested that noise decreased the listeners' certainty in the auditory input, leading them to consider other phonological alternatives, or as McQueen and Huettig, (2012) suggested, the presence of noise changes the perceptual weight assigned to acoustic information. That is, it is possible that the interaction of working memory load and memory span found in our study may reflect an increase in the uncertainty in the auditory input. It is possible that these differences will not be easily detected when tested in ideal listening conditions.

Our results might be interpreted in light of the ease of language understanding model (Rönnberg et al., 2013) that describes the way in which working memory is involved in spoken language understanding. These authors suggest that individuals with high working memory capacity can deploy more resources to different aspects of the speech perception task. According to this model, a greater mismatch between sensory and mental representations, and hence a higher involvement of working memory processes, is predicted with less favorable speech-to-noise ratios. Following the ease of language understanding model, it can be predicted that under adverse listening conditions, individuals with higher working memory capacity will better adapt to task demands than individuals with lower working memory capacity. In the current study, this prediction would be reflected by the difference in the timeline for target fixations between participants with lower and higher memory span, under a high working memory load condition. Individuals with lower memory span were less able to adapt to the increased working memory load and therefore were less efficient in discriminating the target spoken word from its phonological competitor. The complementary Framework for Understanding Effortful Listening (Pichora-Fuller et al., 2016) suggests that speech processing could be impacted by individual differences in maximum resource capacity, especially in increased perceptual effort conditions such as presence of background noise and working memory load.

One should also entertain the possibility that executive functions, other than working memory span, might underlay some of the effects noted in our study (Wingfield, 2016). For example, effectiveness of attentional switch was found to correlate with the perception of fundamental speech contrasts (Ou & Law, 2017). In their studies, Ou and Law suggest that the perception of tonal differences in Cantonese were related to scores on an attentional switch task. It is possible that in our study, participants were switching their attention between the two tasks (spoken word processing and digit memorization) instead of processing them in tandem, and that this switch had an impact on the observed effects.

Caveat and Future Studies

We note that we did not control for musical training background that some argue may affect speech processing (Parbery-Clark, Skoe, Lam, & Kraus, 2009). In addition, hearing was tested using pure-tone audiometric thresholds that do not necessarily eliminate the presence of so-called hidden hearing loss (Barbee et al., 2018). It is reasonable to assume that these factors are randomly distributed across groups. However, it is possible that the percentage of individuals with musical training experience might be greater in the higher memory span group than the lower memory span group (see George & Coch, 2011). Future studies may wish to examine this in a dedicated study testing the possible effects of musical training and hidden hearing loss. Future studies may also wish to examine this paradigm tailoring intensity to individuals' pure tone and speech reception thresholds (e.g., see Ben-David, Avivi-Reich, & Schneider, 2016).

Summary

The present results show that differences in memory span have an effect on the timeline for spoken word recognition in adverse conditions, as the word unfolds. Specifically, when a four-digit working memory load was imposed, listeners with lower memory span were delayed in discriminating target from sound sharing competitor by ~550 ms, relative to higher span listeners. However, no effect was evident in accuracy for word recognition. Our results can be taken as further support for the use of eye-tracking as a sensitive measure of online speech processing. Finally, it is notable that if we were to collect the data in the current study without taking into account the intrinsic differences in memory capacity, the effect of extrinsic working memory load would significantly diminish. This highlights the importance of noting individual differences in cognitive capacity when testing speech processing, even in normal hearing young adults.

Appendix A. R Packages Used for Growth Curve Analysis

Software versions: (RStudio & R were acquired via Anaconda Navigator, which was acquired from the anaconda website: <https://www.anaconda.com/>)

Anaconda Navigator	1.8.7
RStudio	1.1.423
R	3.4.3

R Packages (all acquired from CRAN <https://cran.r-project.org/>)

Packages used in prepping the data and scripting:

VWPre	1.1.0
dplyr	0.7.6
forcats	0.2.0
knitr	1.18
magrittr	1.5
purrr	0.2.4
readr	1.1.1
stringr	1.2.0
tibble	1.4.2
tidyr	0.7.2
tidyverse	1.2.1
zeallot	0.1.0

Packages used in modeling:

lme4	1.1–15
Hmisc	4.1–1
nloptr	1.0.4

Full code available at: <https://github.com/CANlab-IDC/differences-wm-capacity-analysis-2019>

Acknowledgments

The authors wish to thank Dalith Tal-Shir, Joshua E. Skrzypek, and Juliet Gavison for their work on this project.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: A. W. research is supported by the U.S. National Institutes of Health under award number R01 AG019714.

Supplemental Material

Supplemental material is available for this article online.

References

- 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. *International Journal of Audiology, 47*, 53–71. doi:10.1080/14992020802301142
- Alloppenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language, 38*(4), 419–439. doi:10.1006/jmla.1997.2558

- Arnold, J. E., Fagnano, M., & Tanenhaus, M. K. (2003). Disfluencies signal thee, um, new information. *Journal of Psycholinguistic Research*, 32(1), 25–36. doi:10.1023/A:1021980931292
- Ayasse, N. D., Lash, A., & Wingfield, A. (2017). Effort not speed characterizes comprehension of spoken sentences by older adults with mild hearing impairment. *Frontiers in Aging Neuroscience*, 8, 329. doi:10.3389/fnagi.2016.00329
- Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology*, 63, 1–29. doi:10.1146/annurev-psych-120710-100422
- Baddeley, A. D., & Hitch, G. (1974). Working memory. *Psychology of Learning and Motivation*, 8, 47–89. doi:10.1016/s0079-7421(08)60452-1
- Barbee, C., James, J., Park, J., Smith, E. M., Johnson, C. E., Clifton, S., & Danhauer, J. L. (2018). Effectiveness of auditory measures for detecting hidden hearing loss and/or cochlear synaptopathy: A systematic review. *Seminars in Hearing*, 39(02), 172–209. doi:10.1055/s-0038-1641743
- Ben-David, B. M., Avivi-Reich, M., & Schneider, B. A. (2016). Does the degree of linguistic experience (native versus non-native) modulate the degree to which listeners can benefit from a delay between the onset of the maskers and the onset of the target speech? *Hearing Research*, 341, 9–18. doi:10.1016/j.heares.2016.07.016
- Ben-David, B. M., Chambers, C. G., Daneman, M., Pichora-Fuller, M. K., Reingold, E. M., & Schneider, B. A. (2011). Effects of aging and noise on real-time spoken word recognition: Evidence from eye movements. *Journal of Speech, Language, and Hearing Research*, 54(1), 243–262. doi:10.1044/1092-4388(2010/09-0233)
- Ben-David, B. M., Erel, H., Goy, H., & Schneider, B. A. (2015). “Older is always better”: Age-related differences in vocabulary scores across 16 years. *Psychology and Aging*, 30(4), 856. doi: 10.1037/pag0000051
- Ben-David, B. M., Tse, Y. Y., & Schneider, B. A. (2012). Does it take older adults longer than younger adults to perceptually segregate a speech target from a background masker? *Hearing Research*, 290(1–2), 55–63. doi:10.1016/j.heares.2012.04.022
- Benichov, J., Cox, L. C., Tun, P. A., & Wingfield, A. (2012). Word recognition within a linguistic context: Effects of age, hearing acuity, verbal ability and cognitive function. *Ear and Hearing*, 32(2), 250. doi:10.1097/aud.0b013e31822f680f
- Besser, J., Koelewijn, T., Zekveld, A. A., Kramer, S. E., & Festen, J. M. (2013). How linguistic closure and verbal working memory relate to speech recognition in noise—A review. *Trends in Amplification*, 17(2), 75–93. doi: 10.1177/1084713813495459
- Brouwer, S., & Bradlow, A. R. (2016). The temporal dynamics of spoken word recognition in adverse listening conditions. *Journal of Psycholinguistic Research*, 45(5), 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. *Psychonomic Bulletin & Review*, 16(5), 893–900. doi:10.3758/PBR.16.5.893
- Cooper, R. M. (1974). The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology*, 6, 84–107. doi: 10.1016/0010-0285(74)90005-x
- Cowan, N. (1999). An embedded-process model of working memory. In A. Miyake, & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 62–101). Cambridge, England: Cambridge University Press. doi: 10.1017/cbo9781139174909.006
- Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review*, 3(4), 422–433. doi: 10.3758/bf03214546
- Dryden, A., Allen, H. A., Henshaw, H., & Heinrich, A. (2017). The association between cognitive performance and speech-in-noise perception for adult listeners: A systematic literature review and meta-analysis. *Trends in Hearing*, 21, 2331216517744675. doi:10.1177/2331216517744675
- Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, 11(1), 19–23. doi:10.1111/1467-8721.00160
- Ezzatian, P., Avivi, M., & Schneider, B. A. (2010). Do nonnative listeners benefit as much as native listeners from spatial cues that release speech from masking? *Speech Communication*, 52(11–12), 919–929. doi:10.1016/j.specom.2010.04.001
- Füllgrabe, C., & Rosen, S. (2016). On the (Un)importance of working memory in speech-in-noise processing for listeners with normal hearing thresholds. *Frontiers in Psychology*, 7, 1268. doi:10.3389/fpsyg.2016.01268
- George, E. M., & Coch, D. (2011). Music training and working memory: An ERP study. *Neuropsychologia*, 49(5), 1083–1094. doi:10.1016/j.neuropsychologia.2011.02.001
- Goodman, L. (2001). Translation of WAIS-III - Wechsler Adult Intelligence Scale. Jerusalem, Israel: Psych tech.
- Gordon-Salant, S., & Cole, S. S. (2016). Effects of age and working memory capacity on speech recognition performance in noise among listeners with normal hearing. *Ear and Hearing*, 37(5), 593–602. doi:10.1097/aud.0000000000000316
- Hadar, B., Skrzypek, J. E., Wingfield, A., & Ben-David, B. M. (2016). Working memory load affects processing time in spoken word recognition: Evidence from eye-movements. *Frontiers in Neuroscience*, 10, 221. doi:10.3389/fnins.2016.00221
- Huetting, F., & Janse, E. (2016). Individual differences in working memory and processing speed predict anticipatory spoken language processing in the visual world. *Language, Cognition and Neuroscience*, 31(1), 80–93. doi:10.1080/23273798.2015.1047459
- Huetting, F., Rommers, J., & Meyer, A. S. (2011). Using the visual world paradigm to study language processing: A review and critical evaluation. *Acta Psychologica*, 137(2), 151–171. doi:10.1016/j.actpsy.2010.11.003
- Kaiser, E., & Trueswell, J. C. (2008). Interpreting pronouns and demonstratives in Finnish: Evidence for a form-specific approach to reference resolution. *Language and Cognitive Processes*, 23(5), 709–748. doi:10.1080/01690960701771220
- Lash, A., & Wingfield, A. (2014). A Bruner-Potter effect in audition? Spoken word recognition in adult aging. *Psychology and Aging*, 29(4), 907. doi:10.1037/a0037829
- Lezak, M. D. (1995). *Neuropsychological assessment* (3rd ed.). New York, NY: Oxford University Press.

- Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing, 19*(1), 1. doi:10.1097/00003446-199802000-00001
- Marslen-Wilson, W. D. (1990). Activation, competition, and frequency in lexical access. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing* (pp. 148–172). Cambridge, England: MIT Press. doi:10.1017/s0142716400005798
- McCabe, D. P., Roediger, H. L., McDaniel, M. A., Balota, D. A., & Hambrick, D. Z. (2010). The relationship between working memory capacity and executive functioning: Evidence for a common executive attention construct. *Neuropsychology, 24*(2), 222. doi:10.1037/a0017619
- McQueen, J. M., & Huettig, F. (2012). Changing only the probability that spoken words will be distorted changes how they are recognized. *The Journal of the Acoustical Society of America, 131*(1), 509–517. doi:10.1121/1.3664087
- Mirman, D. (2014). *Growth curve analysis and visualization using R*. Boca Raton, FL: Chapman and Hall/CRC Press.
- Mirman, D., Dixon, J. A., & Magnuson, J. S. (2008). Statistical and computational models of the visual world paradigm: Growth curves and individual differences. *Journal of Memory and Language, 59*(4), 475–494. doi:10.1016/j.jml.2007.11.006
- Oberauer, K. (2002). Access to information in working memory: Exploring the focus of attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*(3), 411. doi:10.1037/0278-7393.28.3.411
- Ou, J., & Law, S.-P. (2017). Cognitive basis of individual differences in speech perception, production and representations: The role of domain general attentional switching. *Attention, Perception, & Psychophysics, 79*(3), 945–963. doi:10.3758/s13414-017-1283-z
- Parbery-Clark, A., Skoe, E., Lam, C., & Kraus, N. (2009). Musician enhancement for speech-in-noise. *Ear and Hearing, 30*(6), 653–661. doi:10.1097/AUD.0b013e3181b412e9
- Pichora-Fuller, M. K., Kramer, S. E., Eckert, M. A., Edwards, B., Hornsby, B. W., Humes, L. E., . . . Wingfield, A. (2016). Hearing impairment and cognitive energy: The Framework for Understanding Effortful Listening (FUEL). *Ear and Hearing, 37*, 5S–27S. doi:10.1097/aud.0000000000000312
- Postle, B. R. (2006). Working memory as an emergent property of the mind and brain. *Neuroscience, 139*(1), 23–38. doi:10.1016/j.neuroscience.2005.06.005
- Rönnberg, J., Rudner, M., Foo, C., & Lunner, T. (2008). Cognition counts: A working memory system for ease of language understanding (ELU). *International Journal of Audiology, 47*(Suppl 2): S99–S105. doi:10.1080/14992020802301167
- Rönnberg, J., Lunner, T., Zekveld, A., Sörqvist, P., Danielsson, H., Lyxell, B., . . . Rudner, M. (2013). The Ease of Language Understanding (ELU) model: Theoretical, empirical, and clinical advances. *Frontiers in Systems Neuroscience, 7*, 31. doi:10.3389/fnsys.2013.00031
- Rossion, B., & Pourtois, G. (2004). Revisiting Snodgrass and Vanderwart's object pictorial set: The role of surface detail in basic-level object recognition. *Perception, 33*(2), 217–236. doi: 10.1068/p5117
- Rudner, M., Lunner, T., Behrens, T., Thorén, E. S., & Rönnberg, J. (2012). Working memory capacity may influence perceived effort during aided speech recognition in noise. *Journal of the American Academy of Audiology, 23*(8), 577–589. doi:10.3766/jaaa.23.7.7
- Sommers, M. S., & Amano, S. (1998). Lexical competition in spoken word recognition by younger and older adults: A comparison of the rime cognate, neighborhood, and cohort. *Journal of Acoustical Society of America, 103*, 2984–2984. doi: 10.1121/1.421676
- Sörqvist, P., & Rönnberg, J. (2012). Episodic long-term memory of spoken discourse masked by speech: What is the role for working memory capacity? *Journal of Speech, Language, and Hearing Research, 55*(1), 210–218. doi:10.1044/1092-4388(2011/10-0353)
- Stenbäck, V., Hällgren, M., Lyxell, B., & Larsby, B. (2015). The Swedish Hayling task, and its relation to working memory, verbal ability, and speech-recognition-in-noise. *Scandinavian Journal of Psychology, 56*(3), 264–272. doi:10.1111/sjop.12206
- Tanenhaus, M. K., Magnuson, J. S., Dahan, D., & Chambers, C. (2000). Eye movements and lexical access in spoken-language comprehension: Evaluating a linking hypothesis between fixations and linguistic processing. *Journal of Psycholinguistic Research, 29*(6), 557–580. doi:10.1023/a:1026464108329
- Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science, 268*(5217), 1632. doi:10.1126/science.7777863
- Wayland, S. C., Wingfield, A., & Goodglass, H. (1989). Recognition of isolated words: The dynamics of cohort reduction. *Applied Psycholinguistics, 10*(4), 475–487. doi:10.1017/s0142716400009048
- Wingfield, A. (2016). Evolution of models of working memory and cognitive resources. *Ear and Hearing, 37*, 35S–43S. doi: 10.1097/AUD.0000000000000310
- Wingfield, A., Goodglass, H., & Lindfield, K. C. (1997). Word recognition from acoustic onsets and acoustic offsets: Effects of cohort size and syllabic stress. *Applied Psycholinguistics, 18*(1), 85–100. doi:10.1017/s0142716400009887
- Wingfield, A., & Tun, P. A. (2007). Cognitive supports and cognitive constraints on comprehension of spoken language. *Journal of the American Academy of Audiology, 18*(7), 548–558. doi:10.3766/jaaa.18.7.3
- Zhang, X., & Samuel, A. G. (2018). Is speech recognition automatic? Lexical competition, but not initial lexical access, requires cognitive resources. *Journal of Memory and Language, 100*, 32–50. doi:10.1016/j.jml.2018.01.002