Cognition 225 (2022) 105153

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

Capacity limits in sentence comprehension: Evidence from dual-task judgements and event-related potentials

Trevor Brothers *

Tufts University, Department of Psychology, United States of America

ARTICLE INFO

Keywords:
Word recognition
Visual attention
Sentence processing
N400
Sentence superiority effect

ABSTRACT

There is an ongoing controversy over whether readers can access the meaning of multiple words, simultaneously. To date, different experimental methods have generated seemingly contradictory evidence in support of serial or parallel processing accounts. For example, dual-task studies suggest that readers can process a maximum of one word at a time (White, Palmer & Boynton, 2018), while ERP studies have demonstrated neural priming effects that are more consistent with parallel activation (Wen, Snell & Grainger, 2019). To help reconcile these views, I measured neural responses and behavioral accuracy in a dual-task sentence comprehension paradigm. Participants saw masked sentences and two-word phrases and had to judge whether or not they were grammatical. Grammatically correct sentences (*This girl is neat*) produced smaller N400 responses compared to scrambled sentences (*Those girl is fled*): an N400 *sentence superiority effect*. Critically, participants' grammaticality judgements on the same trials showed striking capacity limitations, with dual-task deficits closely matching the predictions of a serial, *all-or-none* processing account. Together, these findings suggest that the N400 sentence superiority effect is fully compatible with serial word recognition, and that readers are unable to process multiple sentence positions simultaneously.

1. Introduction

To a novice reader, a page of text may appear as a cluttered and meaningless visual scene. Only by applying selective spatial attention can comprehenders extract the meanings of individual words and sentences. While reader's eye-movements through a text are serial – with only one word fixated at a time – it is unclear whether similar processing constraints apply within the visual word recognition system. Specifically, it is unclear if word recognition also occurs *serially*, or if all words surrounding fixation are processed in *parallel*, without competing for attentional resources. In this study, I tested these competing accounts by examining the attentional capacity limits of the visual word recognition system (Duncan, 1980).

1.1. Evidence for serial procession

When viewing crowed arrays, certain visual features are thought to be processed in parallel, with apparently unlimited capacity (Treisman, 1986; Treisman & Gelade, 1980). For example, objects with a unique color "pop-out" from the visual field, allowing them to be detected quickly regardless of the number of surrounding distractors (see Fig. 1). In contrast, searching for words in cluttered arrays is remarkably inefficient, with additional distractor words producing a linear increase in search times (50–150 ms per item, Flowers & Lohr, 1985; Karlin & Bower, 1976; Harris, Pashler, & Coburn, 2004). These dissociations have led some researchers to conclude that "word search never suggests parallel processing" (Duncan, 1987), and that "it is not possible to read two messages in two parts of the [visual] field at the same time" (Wolfe, 1994).

Further evidence of capacity limits in visual word recognition comes from dual-task paradigms (Harris et al., 2004; Reichle, Vanyukov, Laurent, & Warren, 2008; Scharff, Palmer, & Moore, 2011; White, Palmer, Boynton, & Yeatman, 2019). For example, in a recent study (White, Palmer, & Boynton, 2018) two masked words were presented briefly to the left and right of fixation as participants monitored for a specific semantic category (e.g. occupations). Across blocks, participants were told to attend either to a single word (*single-task condition*) or to *both* words simultaneously (*dual-task condition*). Recognition accuracy was severely impaired in the dual-task condition, and the magnitude of this dual-task deficit was consistent with "all-or-none" processing of a

* Corresponding author at: Tufts University, Department of Psychology, 419 Boston Ave., Medford, MA 02155, United States of America.
E-mail address: trevor.brothers@gmail.com.

https://doi.org/10.1016/j.cognition.2022.105153
Received 1 April 2021; Received in revised form 15 April 2022; Accepted 26 April 2022
Available online 9 May 2022

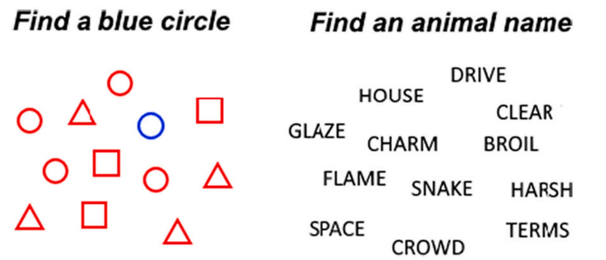

Fig. 1. A comparison of efficient visual search for color information (left) and inefficient search for lexical information (right).

single word on each trial. In contrast, when participants monitored the same stimuli for changes in font color, participants showed minimal dual-task deficits. These findings suggest that the low-level perceptual features (e.g. color) are processing in parallel across the visual field, while word recognition depends a limited-capacity, serial attention mechanism (for similar findings, see White, Palmer, & Boynton, 2020).

1.2. Evidence for parallel processing

While these findings are consistent with serial lexical processing, this position has not been universally adopted. For example, Snell and Grainger (2019) have proposed that readers are actually parallel processers, based on studies using simultaneously presented, multi-word arrays. In these experiments, participants make categorical judgements about a centrally presented word, while surrounding 'flanker words' were presented in close spatial proximity. When flankers are congruent (e.g. from the same category), this improved the speed and accuracy of behavioral responses (Snell, Meeter, & Grainger, 2017). While word flanker effects of this kind have a long history (Bradshaw, 1974; Dallas & Merikle, 1976; Shaffer & LaBerge, 1979; Underwood, 1976, 1981; Underwood & Thwaites, 1982), it is unclear if these findings provide direct evidence for simultaneous lexical processing. Specifically, with longer stimulus durations or inadequate masking, readers could process target and flanker words *sequentially*, which could also influence categorization behavior (see White et al., 2018, 2020; Broadbent & Gathercole, 1990 for similar arguments).

Snell and Grainger (2019) also cite another classic phenomenon: the *sentence superiority* effect (Cattell, 1886; Juola, Ward, & McNamara, 1982) as evidence of parallel lexical processing. In these studies, words are recognized more accurately in congruent sentence contexts (*The man can run*) than in ungrammatical contexts (*Run man can the*) (Jordan & Thomas, 2002; Marslen-Wilson & Tyler, 1980), and these effects occur even when sentences are flashed simultaneously and then masked (Asano & Yokosawa, 2011; Snell & Grainger, 2017). Most surprisingly, in a recent event-related potential study (Wen, Snell, & Grainger, 2019), simultaneously presented 'intact sentences' produced both improved word identification *and* a reduction in the amplitude of the N400 response, which is thought to reflect the difficulty of lexico-semantic retrieval (Kutas & Federmeier, 2011). Given the early time-course of these ERP differences (300-500 ms), the authors suggested that this effect may provide an online neural marker for the "parallel processing of word identities" (Wen et al., 2019).

To summarize, two recent studies have provided apparently contradictory evidence for serial vs. parallel lexical processing, using a dual-task paradigm (White et al., 2018) and simultaneous sentence presentation (Wen et al., 2019). What might account for these discrepancies? One possibility is that all word recognition involves a serial attentional bottleneck, but that the sentence superiority effect is actually compatible with serial recognition. As discussed above, with longer stimulus durations (e.g. 200 ms) readers may have enough time to process multiple words *sequentially* before the appearance of a mask. Under this account, the initial stages of lexical processing would occur serially, with the semantic and syntactic features of each word being stored in working memory. Consequently, in intact contexts (*The man can run*) an initially attended word ("can") could provide syntactic constraints to prime the recognition of subsequently attended words ("run"). If this is the case, simultaneously presented intact sentences could produce neural and behavioral facilitation, even within a serial processing system.

1.3. The current study

In the present study, I tested this possibility by recording behavioral and ERP responses simultaneously in a dual-task sentence comprehension paradigm. On single-task trials, participants saw a masked, two-word phrase (*Those girl*…), and were asked to judge whether or not it was grammatical. On dual-task trials participants saw a masked 4-word sentence (*Those girl is neat*), and were asked to provide two separate grammaticality judgements, one for the first two words of the sentence, and one for the last two words. Using ERPs, I compared brain responses to fully grammatical ("intact") sentences and fully ungrammatical ("scrambled") sentences, in an attempt to replicate the N400 sentence superiority effect. Critically, at the same time, I compared single-task and dual-task performance using Attention Operating Characteristics (AOCs), to quantify the attentional capacity limits of readers during sentence processing (Sperling & Melchner, 1978; White et al., 2018).

According to serial models, readers use spatial attention to bind letters into a holistic lexical representation. To avoid interference, this spotlight of attention encompasses only a single word at a time (Reichle, Rayner, & Pollatsek, 2003), which places fundamental limits on sentence processing rates. In contrast, in parallel models, readers deploy a broad *gradient* of attention, and any words falling within this gradient are processed at a fixed rate, dependent on their distance from fixation (Engbert, Nuthmann, Richter, & Kliegl, 2005; Snell, 2018). As a consequence, these two models make very different predictions regarding how the *number* of simultaneously presented words will influence identification accuracy (Reichle et al., 2008). Specifically, in a dual-task paradigm, parallel attention models predict little to no dual-task deficit, while serial models predict approximately a 50% reduction in sensitivity.

To preview our results, ERP recordings showed a robust N400 sentence superiority effect, consistent with the results of Wen et al. (2019). Critically, participants' grammaticality judgements on the same trials showed striking capacity limitations, with dual-task deficits closely matching the predictions of a serial, *all-or-none* processing model. Considered together, these findings prompt a critical re-evaluation of the sentence superiority effect, while also providing a difficult empirical challenge for parallel processing models of lexical access.

2. Materials and methods

2.1. Participants

For this study I recruited 31 participants from Tufts University and the surrounding community (13 female, mean age = 20 years). All participants were right-handed as assessed by the Edinburgh Handedness Inventory with normal or corrected-to-normal vision. They were also all native English speakers with no exposure to other languages before the age of five. Participants provided informed consent and were compensated for their time. In addition, all protocols were approved by Tufts University Social, Behavioral, and Educational Research Institutional Review Board. EEG data from one participant was excluded due to excessive artifact.

2.2. Materials

For this experiment, I generated 192 short, declarative sentences (*"Those kids had fled."*). Each sentence contained four words, 1 to 5 letters in length. Words in the first-half and the second-half of the sentence were matched in length (3.6 characters, SD = 1.1) and lexical frequency (Log per billion = 5.6, SD = 1.2, Brysbaert & New, 2009). Including spaces and punctuation, sentences had an average length of 18.6 characters (SD = 1.6).

In prior studies examining the sentence superiority effect (Snell & Grainger, 2017; Wen et al., 2019) words were typically shuffled *within* sentences. Here, to avoid word-position confounds, I instead adopted a fully counterbalanced design by swapping sentence-initial or sentence-final words across sentence pairs (see Table 1). These word-swaps resulted in one of three types of grammatical errors: 1) number agreement errors (e.g. *she steal*…; *we moves*…), 2) determiner agreement errors (e.g. *an song*…; *a award*…), or 3) syntactic category errors (e.g. *her bring*…; *don't hair*…). Because words always appeared in the same sentence position across conditions, participants were unable to make grammaticality judgements based on the position of a single word (e.g. a sentence ending in the word "…*an.*").

Grammatical errors could appear in the first half of the sentence, the second half of the sentence, or both. Therefore, four versions of each sentence were created, containing zero, one, or two grammatical errors (Grammatical, Left Error, Right Error, Ungrammatical). To generate items for the single-task condition, one third of the grammatical and ungrammatical sentences on each list were split into two-word phrases ("those kids…", "…had fled."), which were presented on separate trials.

At the beginning of each session, participants were assigned to one of 12 experimental lists. This counterbalancing scheme ensured that each item appeared equally often in the four experimental conditions (Grammatical, Right Error, Left Error, Ungrammatical) and that items were equally likely to appear on single-task and dual-task trials. Sentences were never repeated within an experimental session.

2.3. Methods

During the experiment, participants were seated comfortably in a sound-attenuated room, and stimuli were presented on a computer monitor at a distance of 1.5 m. All words were presented in white Consolas font on a black background, with three characters subtending one degree of visual angle. On dual-task trials, a complete, four-word sentence was presented centrally for 200 ms, followed by a random letter mask (*jxpdibfhq*>> > *smjosyrva*). The center of this mask contained a visual prompt (<<< or >>>) indicating which side of the sentence required the first grammaticality judgement. The prompt remained on the screen until the participants provided a response. Participants were then prompted to provide a second grammaticality judgement for the other half of the sentence. The order of these prompts (left vs. right) was randomized across trials (see Fig. 2), and no feedback was provided.

On single-task trials, only two words appeared on the screen (left or right), with all other characters replaced by slashes ("*a song /// /////.*"). Single-response items also appeared for 200 ms, followed by a random letter mask, and participants judged whether or not this two-word phrase contained a grammatical error. After a response was recorded, the letter mask disappeared, and the next trial began after a jittered inter-stimulus interval (1600–2400 ms). All stimuli were presented in lowercase, fixed width font, which ensured that words in the single-task and dual-task trials appeared in the same spatial positions.

At the beginning of the experimental session, participants were told that they would read a series of sentences and two-word phrases and judge whether or not they contained grammatical errors. Participants were given several examples of each error type, followed by a series of practice trials. Participants were told that, on dual-task trials, the presence of a grammatical error on one side of the sentence was not diagnostic of the presence (or absence) of a grammatical error on the other side. Participants were told to stress accuracy rather than speed, and all responses were untimed.

Throughout the experiment, participants were told to fix their eyes in the center of the screen between two vertical lines, which remained in place throughout the experiment (see Fig. 2). Eye movements were monitored throughout the experiment using electo-oculograms attached to the outer canthus of each eye. In order to minimize differences in arousal or pre-stimulus attentional allocation across conditions, single-task and dual-task trials were presented in a randomized order, in the same block. After an initial practice session, participants saw 256 trials, half in the single-task condition and half in the dual-task condition. The entire experiment lasted approximately 30 min, with a short break provided at the halfway point.

Throughout the experiment, EEG signals were recorded continuously from the scalp using 32 active electrodes in a modified 10–20 montage (Biosemi Active-Two). Signals were digitized at 512 Hz with a passband of DC - 102 Hz. The EEG was referenced offline to the average of the right and left mastoids. A 0.1 to 30 Hz bandpass filter was then applied, and EEG responses were segmented into epochs (−200 ms to 1000 ms), time-locked to the onset of each sentence. EEG activity due to blinks was removed using independent components analysis, and any epochs containing residual artifact or voltage deflections greater than ±75 μV were rejected prior to analysis (3.3% of trials). There were no significant differences in artifact rejection rates across conditions ($F < 1$).

2.4. Analyses

2.4.1. Attention operating characteristics (AOC)

Using AOC functions, I compared participants' performance in single- and dual-task conditions to the predictions of different models of attentional allocation (Sperling & Melchner, 1978; White et al., 2018). In these analyses, accuracy values in the two single-response conditions (left and right) are plotted along each axis, and accuracy in the dual-task condition is plotted as a single point in two-dimensional space (see Fig. 3). Different models of attentional make different predictions regarding the shape this AOC function:

1) According to an *unlimited capacity, parallel-processing* account, readers can extract information from multiple word positions in parallel, with no competition for attentional resources. Because the same amount of information is extracted from each sentence position on both single-response and dual-task trials, this model predicts no dual-task deficits. Therefore, in Fig. 3, dual-task performance should fall on the "independence point" at the intersection of the two dotted lines.
2) According to a *fixed capacity, serial processing* account, readers can process a maximum of one word at a time, with attention shifts only occurring after this word has been successfully identified. For briefly presented, masked stimuli (e.g. 200 ms), participants will only have time to process a maximum of two words. Therefore, when asked to make a grammaticality judgement for an unattended location, participants will simply guess, resulting in chance-level accuracy (50%). Because the dual-task condition will contain a mixture of *attend-left* and *attend-right* trials, dual-task performance will fall somewhere along the black diagonal line (see Fig. 3).
3) Finally, a *fixed capacity, parallel processing* model predicts dual-task performance between these two extremes. According to this model, readers' attentional capacities are limited, but attention can be divided among all four word-positions within a single trial (due to parallel processing or rapid attentional switching; Shaw, 1978). As readers allocate different attentional strengths to the left and right

Table 1
Example sentences and counterbalancing scheme.

Sentence Type	Item 1	Item 2
Grammatical (good/good)	*Those kids had fled.*	*This girl is neat.*
Right Error (good/bad)	*Those kids had neat.*	*This girl is fled.*
Left Error (bad/good)	*This kids had fled.*	*Those girl is neat.*
Ungrammatical (bad/bad)	*This kids had neat.*	*Those girl is fled.*

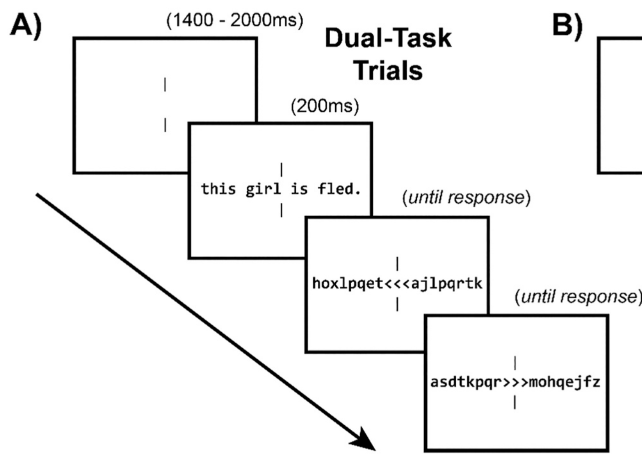

Fig. 2. A) On the dual-task trials, participants provided two grammaticality judgements, one for the first two words of the sentence ("this girl…") and one for the final two words ("…is fled."). Response order (left/right) was randomized across trials. B) On the single-task trials, only the first or last two words of the sentence were presented, and participants provided a single grammaticality judgement.

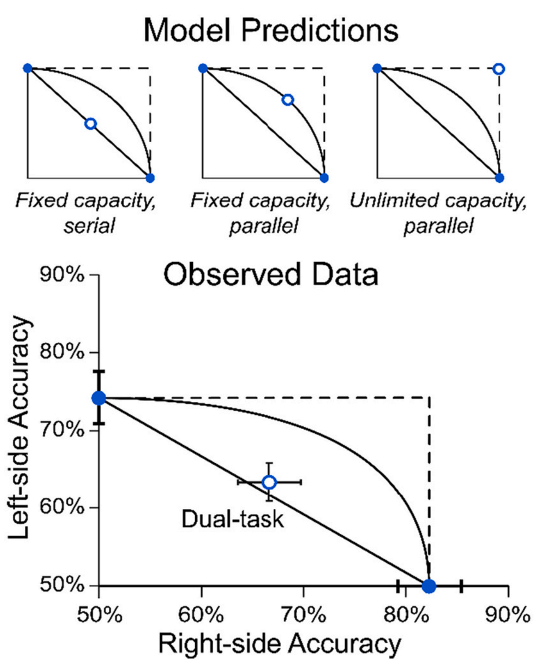

Fig. 3. Mean Attention Operating Characteristics (AOC) in the grammaticality judgement task. Solid points on the x- and y-axes reflect behavioral accuracy on single-task trials (percent correct), and the open point reflects accuracy on dual-task trials. The upper panel shows predictions for three classes of visual attention models. The observed data (lower panel) was consistent with a *fixed capacity, serial processing* model. Error-bars show 95% confidence intervals.

sentence positions, the predictions of the *fixed capacity parallel model* trace a curve through the center of the AOC plot (White et al., 2018). These model predictions were calculated using the equations described in Bonnel and Haftser (1998).

2.4.2. Event-related potentials

Event-related potentials were analyzed using cluster-mass permutation tests, examining ERP differences from 100 ms to 1000 ms across 29 electrode scalp sites (Groppe, Urbach, & Kutas, 2011). I used a cluster-forming threshold of $p < .05$, and condition labels were randomly shuffled across 10,000 permutations to construct null distributions. Finally, I also performed a set of planned contrasts in a spatio-temporal region of interest (300-500 ms; Fz, AF3/4, F3/4, FC1/2), where N400 sentence superiority effects were reported previously. These pairwise contrasts allowed me to test for graded grammaticality effects on the N400 (Grammatical < Single-error < Ungrammatical). However, it should be noted that any graded effect would not be necessarily diagnostic of either parallel or serial processing (see Discussion section).

3. Results

3.1. Behavioral results

Across all trials, participants were more accurate when making judgements in the right visual field compared to the left (right = 68.8%; left: 74.5%, z = 3.21, p = .001), consistent with the well-known left hemisphere advantage for language processing (Bradshaw & Nettleton, 1981; White, 1969), see Table 2. More importantly, participants' accuracy was much higher on single-task trials (78.2%, SD = 6.5%) than dual-task trials (65.0%, SD = 7.3%). This effect was highly significant and a dual-task deficit was observed in all 31 participants ($diff_{acc}$ = 13.2% ± 0.9%, z = 14.38, $p < .0001$).[1]

In the dual-task condition, there was no significant difference in accuracy for participants' first and second grammaticality judgements

Table 2
Mean accuracy and signal detection measures across conditions (with standard deviations).

	Single-Task Condition		Dual-Task Condition	
	Left side	Right side	Left side	Right side
Proportion Correct	.74 (.10)	.82 (.09)	.63 (.07)	.67 (.09)
Sensitivity (d')	1.46 (0.70)	2.06 (0.64)	0.73 (0.41)	1.04 (0.58)
Bias (c)	0.25 (0.33)	0.40 (0.30)	0.17 (0.21)	0.45 (0.32)

[1] Across all analyses, a similar pattern of results was observed when examining reader's sensitivity to grammatical errors using signal detection theory (d'). In terms of response bias (c), participants showed a slight bias to respond "correct", which was equivalent in the single-task and dual-task conditions (single: c = 0.32; dual: c = 0.31, t(30) = 0.32, p = .75), but larger when judging words in the right visual field (right: c = 0.42, left: c = 0.21, t(30) = 6.24, $p < .001$).

(first: 64.9%, second: 65.2%, z = 0.25, p = .81). This suggests that performance costs on dual-task trials are unlikely to reflect differences in response timing or working memory demands across conditions.[2]

As seen in Fig. 3, performance in the grammaticality judgement task aligned closely with the predictions of a *fixed capacity, serial processing* model. When AOC plots were fit to individual participant data, dual-task performance had a mean distance of 19.8% ± 3.0% from the "independence point" predicted by an unlimited capacity parallel model, t(30) = 12.73, $p < .0001$. Accuracy on dual-task trials also did not match the predictions of the *fixed capacity parallel* model, with a mean distance of 6.6% ± 2.4%, t(30) = 5.32, $p < .0001$. The average distance from the nearest point on the fixed capacity, serial processing line was 1.5% ± 2.2%, which did not differ significantly from zero, t(30) = 1.37, p = .18. Dual-task performance was near the mid-point of this serial processing line, suggesting that participants were equally likely to attend to the left or the right side of the sentence on each trial.[3]

3.2. Compatibility effects

In detection and categorization tasks, participants often show higher accuracies when "compatible" stimuli are presented at nearby spatial locations – matching the target either perceptually or in terms of response requirements (Eriksen & Eriksen, 1974; Simon, 1990). These compatibility effects are often thought to reflect parallel processing of multiple spatial locations or, alternately, errors in attentional selection or filtering (Yantis & Johnston, 1990).

In this analysis, I compared readers' behavioral accuracy on "compatible" trials, where the same response was required for both sentence positions (Grammatical, Ungrammatical), and "incompatible" trials, where different responses were required (Left Error, Right Error). On dual-task trials, response compatibility had no significant effect on accuracy (compatible: 65.4%, SD = 6.4%; incompatible: 64.7% SD = 7.1%, z = 0.46, p = .64), and a Bayesian t-test provided moderate support for the null hypothesis (JZS: BF_{01} = 4.5, scaling factor = 0.707). This absence of cross-talk suggests that participants were unable to extract syntactic information from both sides of the display, prior to the onset of the mask. Again, this is consistent with the predictions of a serial processing account.

3.3. ERP results

To examine the influence of grammatical errors on online neural responses, I analyzed ERPs in the dual-task condition, time-locked to sentence onset. ERPs on dual-task trials showed a series of evoked visual components (P1, N1, P2) followed by a negative peak (N400) that was broadly distributed across the scalp (see Fig. 4). Beginning approximately 300 ms after sentence onset, more negative ERP amplitudes were observed for Grammatical versus Ungrammatical sentences, consistent with the N400 *sentence superiority* effect observed by Wen and colleagues. This negativity was most pronounced over frontal and central electrode sites, and these ERP differences were relatively sustained from 300 ms until the end of the ERP epoch. A cluster-mass permutation test, comparing ERP responses to Ungrammatical and Gramamtical sentences, revealed a significant negative-going cluster over central and frontal electrode sites (314 ms–865 ms, p = .008 peak-electrode: AF3).

In addition, dual-task trials with a single grammatical error (Left Error, Right Error) elicited an intermediate N400 response (see Fig. 5) that did not differ as a function of error position ($t < 1$). To verify this graded N400 difference, I compared ERP responses in a frontal-central cluster of electrode sites from 300 to 500 ms (Grammatical: −0.99 μV; Single-error: −1.62 μV; Ungrammatical: −2.05 μV). A repeated-measures ANOVA within this ROI revealed a robust effect of Condition (F(2,58) = 10.22, $p < .001$). Pair-wise comparisons revealed significant differences for Grammatical vs. Ungrammatical (t(29) = −3.96, $p < .001$) and Grammatical vs. Single-error sentences (t(29) = −2.98, p = .006). The difference between Single-error and Ungrammatical sentences also approached significant (t(29) = −1.92, p = .06).[4]

Finally, similar to the behavioral results reported above, there were no "response compatibility" effects on the amplitude of the N400 (compatible: −1.52 μV, incompatible: −1.62 μV, t(29) = −0.59, p = .56, JZS: BF_{01} = 4.4). This suggests that the N400 is sensitive to syntactic constraints (Ungrammatical > Grammatical), but not to conflicting response requirements.

4. Discussion

In this experiment, I used a dual-task paradigm to explore attentional capacity limits in multi-word, sentence comprehension. Participants saw masked sentences and two-word phrases and judged whether they contained grammatical errors (*This girl*… vs. *Those girl*…). By comparing reader's sensitivity to grammatical violations under single-task and dual-task conditions, I tested whether sentence processing is subject to a serial attentional bottleneck, or if it occurs across multiple word positions in parallel. I also recorded ERP responses simultaneously to determine whether briefly presented sentences would elicit an N400 *sentence superiority* effect (Ungrammatical > Grammatical). In combination with the results of our behavioral analyses, this allowed us to infer whether any observed N400 differences were more consistent with serial or parallel lexical processing.

4.1. Behavioral evidence

Our behavioral findings were quite clear. Grammaticality judgements were less accurate for four-word sentences (dual-task) than two-word phrases (single-task). On dual-task trials, the sensitivity of grammaticality judgements was reduced by 49% (± 7%), which closely matched the 50% reduction predicted by *fixed capacity, serial processing* accounts. These deficits were incompatible with the 0% reduction predicted by *unlimited capacity, parallel processing* accounts, as well as the intermediate deficit predicted by *fixed capacity, parallel processing* models.

These findings are consistent with previous dual-task studies, which showed that readers also engage in serial processing when categorizing unconnected word pairs (*nose – train*; White et al., 2018). This indicates that the same attentional bottleneck applies in both single-word recognition and sentence comprehension. Taken together, these studies suggest that readers require approximately 100 ms of visual input to decode a single high-frequency word (White et al., 2018), and 200 ms to decode a short two-word phrase. They also support the idea that, with brief

[2] To further assess the role of working memory demands, a new group of participants (N = 12) performed the same grammaticality judgement task, but stimulus presentation times were increased from 200 ms to 2000 ms, prior to the onset of the mask. When this early visual processing bottleneck was removed, performance was nearly at ceiling across both tasks (single task: 96%, dual task, 95%). This suggests that differences in working memory demands (one vs. two judgements) cannot explain the large dual-task deficits observed in the main experiment.

[3] Following White et al. (2018) we performed a response contingency analysis, examining whether the accuracy of a participants responses for one side of the sentence depended on their accuracy for the opposite side, within trials. Across participants, this correlation was small and non-significant (mean r = 0.01, SD = 0.08).

[4] Based on the suggestions of a reviewer, we re-analyzed ERP data on Ungrammatical dual-task trials as a function of participant's behavioral responses (correct/incorrect). At frontal electrode sites (300-500 ms) we observed no significant differences due to participants' responses (correct: −1.8 μV, incorrect: −2.2 μV, t(29) = 0.68, p = .50). This suggests that the N400 sentence superiority effect was driven by the grammaticality of each sentence, rather than the participant's behavioral responses.

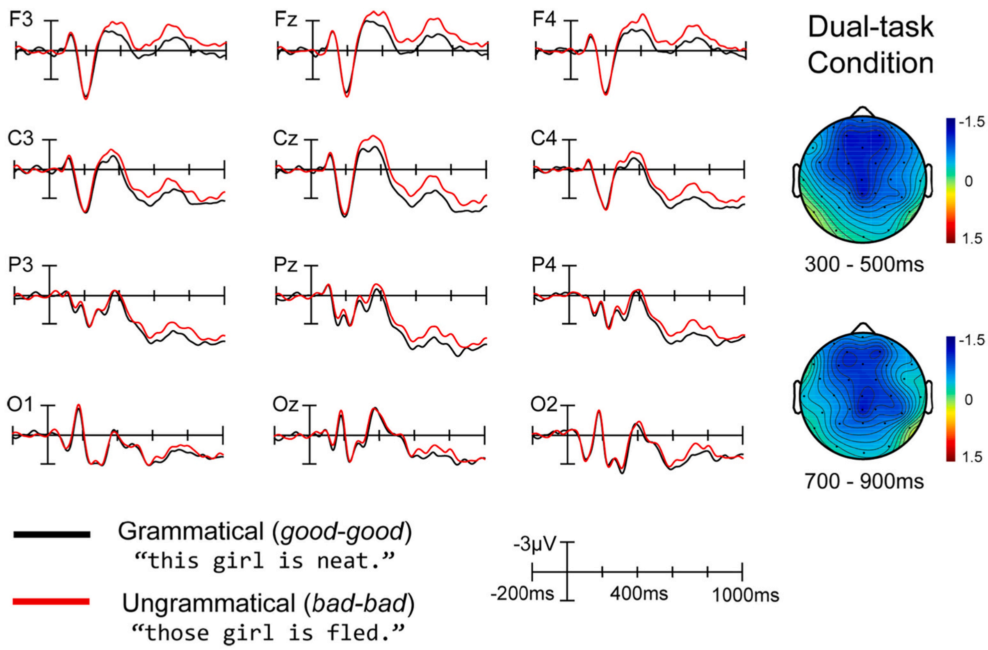

Fig. 4. Grand-average ERPs on dual-task trials, time-locked to the onset of the four-word sentence. The topographic plots (right) show the distribution of the Ungrammatical vs. Grammatical ERP effect in an early (300–500 ms) and late (700–900 ms) time-window.

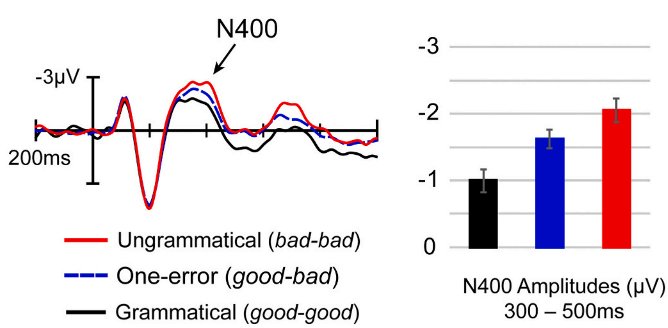

Fig. 5. Grand-average ERPs in the dual-task condition, in a cluster of frontal electrode sites (Fz, AF3/4, F3/4, FC1/2). The bar graph (right) shows mean N400 amplitudes in each condition with ±1 SEM error bars, calculated within-subjects.

stimulus presentations, readers are unable to extract any meaningful information from unattended word positions (Ellis & Marshall, 1978; Inhoff & Topolski, 1992; Pollatsek, Raney, Lagasse, & Rayner, 1993; Williams & Parkin, 1980).[5]

4.2. Theoretical implications

The current findings also provide an important empirical benchmark for computational models of reading (Rayner, 2009). In serial attention models, like *E*-Z Reader (Reichle et al., 2003) comprehension occurs through a series of sequential attention shifts, and readers must complete an early stage of lexical access (the "familiarity check") before they can begin processing the next word of the sentence. These capacity limitations are consistent with the current pattern of behavioral results. On single-task trials ("*a song /// /////.*"), I assume that participants immediately shifted their attention to task relevant portions of the sentence and attempted to process these words, sequentially, before the appearance of the mask. On dual-task trials ("*a song had ended.*") participants selected only one half of the sentence for processing. Due to the brief stimulus presentation, they were unable to extract any additional information from the opposite side, resulting in a robust dual-task deficit.

In contrast, the current findings present a serious challenge to parallel processing frameworks, like SWIFT (Engbert et al., 2005) and OB-1 Reader (Snell, Leipsig, Grainger & Meeter, 2018). In these computational models, readers deploy a broad gradient of attention, centered at the point of fixation. While these models allow some flexibility in the width of this gradient (e.g. across participants or sentences), these models are fundamentally parallel because lexical processing rates remain unchanged, regardless of the number of words falling within this attentional window. Because of this *unlimited-capacity parallel processing* architecture, SWIFT and OB-1 Reader predict *no* dual-task deficits in the current paradigm, which is clearly at odds with our behavioral findings.

Could these parallel models be modified in order to accommodate dual-task deficits? One possibility would be to incorporate a dynamic attentional gradient that allows readers to concentrate attentional resources only on task-relevant subsets of the visual field. In the current task, this would lead to higher accuracy on single-task trials and a slight improvement in model fit. Unfortunately, this *fixed-capacity, parallel* model would still underestimate the dual-task deficits observed in the current paradigm (see Fig. 3). Clearly, additional modelling work will be needed to determine if other factors, like intra-word competition, would be sufficient to reproduce a robust dual-task deficit, without completely abandoning a parallel processing architecture.

Finally, our AOC results also suggest that English readers do not always engage in a strict left-to-right processing strategy. Unlike in natural reading, sentences in this paradigm were presented briefly at fixation, and readers appeared to attend to either the left *or* the right with

[5] It is also worth noting a potential limitation of the current paradigm. In the current study, single-task trials were constructed by replacing task-irrelevant letters with slashes (*This girl // ////*). One benefit of this approach was that it allowed us to randomly intermix *attend left*, *attend right*, and *dual-task* trials within the same block, in order to minimize pre-stimulus differences in attention or arousal. However, with this approach, participants also saw slightly different visual stimuli in the single-task and dual-task conditions, which may have had introduced unintended perceptual effects. In future studies, it will be important to replicate this pattern of dual-task deficits, using a constant visual stimulus and different attentional pre-cues across trials (e.g. see White et al., 2018).

approximately equal frequency across trials. While this reading pattern likely reflects a strategic adaption to the current task, the relaxation of a left-to-right processing strategy may help explain the presence of "right-context" effects in the sentence-superiority paradigm (Snell & Grainger, 2017). Specifically, if readers in this task attend to words sequentially, but out of order, this would explain why the first word of a sentence can also benefit from congruent syntactic constraints.

4.3. ERP evidence

In this experiment, I also recorded ERPs to measure evoked neural responses during simultaneous sentence presentation. On dual-task trials, participants showed more negative ERP responses for Ungrammatical versus Grammatical sentences, particularly at frontal and central electrode sites. Notably, this negativity was very similar to sentence superiority effect reported by Wen et al. (2019), despite clear differences in task requirements (grammaticality judgements vs. cued recall). One could speculate that, rather than a task-specific adaptation, this negativity represents a more general brain response to the processing of grammatical versus ungrammatical strings. However, additional experiments are necessary to confirm this hypothesis.

While the ERP *sentence superiority* effect has been interpreted as a modulation of the N400 response, these effects also differed qualitatively from N400 modulations in typical word-by-word comprehension tasks. For example, while the N400 effects of semantic and syntactic congruity are often observed within 200 ms of word onset (Brothers, Swaab, & Traxler, 2017; Federmeier, Segal, Lombrozo, & Kutas, 2000) in the present study there were no significant ERP differences in the 200 ms to 300 ms time range ($F < 1$). This timing delay is consistent with previous ERP studies that have compared N400 effects under conditions of simultaneous and sequential word presentation (Anderson & Holcomb, 1995; Luka & Van Petten, 2014). In these studies, semantic priming effects were delayed by approximately 100 ms when words pairs were presented simultaneously (300 ms) rather than word-by-word (200 ms). These timing differences suggest that the early stages of lexical access are resource limited, and that the activation of word meanings is delayed when readers are required to shift their attention.

In the current study, even though the visual input was masked 200 ms after sentence onset, participants showed relatively sustained ERP differences, up to 900 ms after sentence onset. This result is reminiscent of an eye-tracking study by Rayner, Liversedge, White, and Vergilino-Perez (2003) who investigated sentence comprehension using a "disappearing text" paradigm. In this study, words disappeared 60 ms after they were first fixated, but readers continued to fixate this blank space for approximately 300 ms before moving on to subsequent regions of the text. More importantly, the duration of these fixations was still influenced by the frequency of the previously fixated word. This suggests that, within 60 ms, readers were able to extract sufficient visual information for word identification, but that lexical and syntactic processing continued, even in the absence of visual inputs (Rayner, Inhoff, Morrison, Slowiaczek, & Bertera, 1981). In future studies, it will be important to determine which sub-stages of word processing show the strongest capacity limitations (e.g. orthographic processing, lexical identification, semantic/syntactic analysis).

Although participants showed clear N400 sentence superiority effects, participants' behavioral results prompted a critical re-evaluation of these ERP differences. Unlike previous authors (Wen et al., 2019), I do not believe the sentence superiority effect provides evidence of simultaneous lexical processing across the visual field. Instead, these findings are fully consistent with serial lexical identification, with initially recognized words providing syntactic or semantic constraints that facilitate the identification of subsequently attended words. If this hypothesis is correct, it makes some clear predictions for future research. Specifically: 1) the onset of the N400 sentence superiority effect should be significantly delayed when comparing simultaneous vs. sequential (word-by-word) presentation, and 2) the sentence superiority effect should be abolished at very short masking durations (<75 ms), when readers are unable to identify more than a single word.

In addition, our results also extended prior ERP findings by demonstrating a graded N400 sentence superiority effect, that tracked the total number of syntactic errors in a sentence (Ungrammatical > Single-error > Grammatical). Again, it is possible to explain this pattern of ERP results by appealing to a serial processing account. If participants were able to attend to only one-half of the sentence on dual-task trials, participants would always encounter a syntactic error in fully ungrammatical sentences, but would encounter an error only 50% of the time on single-error trials (Left Error, Right Error). By averaging trials with attended and unattended syntactic errors, this single-error condition would result in an intermediate N400 response.

5. Conclusion

Selective visual attention is an important component of skilled reading (Casco, Tressoldi, & Dellantonio, 1998; Franceschini, Gori, Ruffino, Pedrolli, & Facoetti, 2012) that also plays a critical role in developmental reading disorders like dyslexia (Valdois, Bosse, & Tainturier, 2004; Vidyasagar & Pammer, 2010). To investigate the role of selective attention during rapid sentence presentation, I investigated readers' attentional capacity limits using behavioral and ERP methods. Consistent with prior ERP findings, I observed clear N400 differences for grammatical and ungrammatical sentences, suggesting that readers use sentence-level constraints to guide word recognition. Critically, behavioral responses on the same trials showed striking capacity limitations, with the magnitude of dual-task deficits supporting a serial "all or none" processing account. Therefore – while it is clear that some simple visual features are processed in parallel across the visual field (e.g. brightness, orientation, color) – the extraction of meaning from abstract symbols appears to require a capacity-limited, serial attention mechanism. These results also provide a cautionary tale for researchers investigating serial and parallel accounts of visual word recognition. Although multi-word arrays can produce a variety of priming effects when *presented* in parallel (Snell & Grainger, 2019), this does not imply that these words are actually *processed* in parallel by the reader. In order to resolve serial versus parallel processing debates, researchers will need to provide evidence from paradigms that can more accurately distinguish these accounts.

Declaration of Competing Interest

I have no conflicts of interest to disclose. This project was funded by National Institute of Child Health and Human Development (R01 HD08252) and the Sidney R. Baer Jr. Foundation. I would like to thank Sophie Greene, Santiago Noriega, and Cassidy Prejean for assistance with data collection, and Gina Kuperberg for her helpful feedback.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cognition.2022.105153.

References

Anderson, J. E., & Holcomb, P. J. (1995). Auditory and visual semantic priming using different stimulus onset asynchronies: An event-related brain potential study. *Psychophysiology, 32*(2), 177–190. https://doi.org/10.1111/j.1469-8986.1995.tb03310.x

Asano, M., & Yokosawa, K. (2011). Rapid extraction of gist from visual text and its influence on word recognition. *The Journal of General Psychology, 138*(2), 127–154. https://doi.org/10.1080/00221309.2010.542510

Bonnel, A.-M., & Haftser, E. R. (1998). Divided attention between simultaneous auditory and visual signals. *Perception & Psychophysics, 60*(2), 179–190. https://doi.org/10.3758/BF03206027

Bradshaw, J. L., & Nettleton, N. C. (1981). The nature of hemispheric specialization in man. *Behavioral and Brain Sciences, 4*(1), 51–63. https://doi.org/10.1017/S0140525X00007548
Bradshaw, J. L. (1974). Peripherally presented and unreported words may bias the perceived meaning of a centrally fixated homograph. *Journal of Experimental Psychology, 103*(6), 1200–1202. https://doi.org/10.1037/h0037371
Broadbent, D. E., & Gathercole, S. E. (1990). The processing of non-target words: Semantic or not? *The Quarterly Journal of Experimental Psychology Section A, 42*(1), 3–37. https://doi.org/10.1080/14640749008401206
Brothers, T., Swaab, T. Y., & Traxler, M. J. (2017). Goals and strategies influence lexical prediction during sentence comprehension. *Journal of Memory and Language, 93*, 203–216. https://doi.org/10.1016/j.jml.2016.10.002
Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods, 41*(4), 977–990. https://doi.org/10.3758/BRM.41.4.977
Casco, C., Tressoldi, P. E., & Dellantonio, A. (1998). Visual selective attention and reading efficiency are related in children. *Cortex, 34*(4), 531–546. https://doi.org/10.1016/S0010-9452(08)70512-4
Cattell, J. M. (1886). The time it takes to see and name objects. *Mind, 11*(41), 63–65.
Dallas, M., & Merikle, P. M. (1976). Semantic processing of non-attended visual information. *Canadian Journal of Psychology/Revue Canadienne de Psychologie, 30*(1), 15–21. https://doi.org/10.1037/h0082040
Duncan, J. (1980). The locus of interference in the perception of simultaneous stimuli. *Psychological Review, 87*(3), 272–300. https://doi.org/10.1037/0033-295X.87.3.272
Duncan, J. (1987). Attention and reading: Wholes and parts in shape recognition: A tutorial review. In *Attention and performance 12: The psychology of reading* (pp. 39–61). Lawrence Erlbaum Associates, Inc.
Ellis, A. W., & Marshall, J. C. (1978). Semantic errors or statistical flukes? A note on allport's "on knowing the meaning of words we are unable to report.". *Quarterly Journal of Experimental Psychology, 30*(3), 569–575. https://doi.org/10.1080/00335557843000142
Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: A dynamical model of saccade generation during Reading. *Psychological Review, 112*(4), 777–813. https://doi.org/10.1037/0033-295X.112.4.777
Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics, 16*(1), 143–149. https://doi.org/10.3758/BF03203267
Federmeier, K. D., Segal, J. B., Lombrozo, T., & Kutas, M. (2000). Brain responses to nouns, verbs and class-ambiguous words in context. *Brain, 123*, 2552–2566.
Flowers, J. H., & Lohr, D. J. (1985). How does familiarity affect visual search for letter strings? *Perception & Psychophysics, 37*(6), 557–567. https://doi.org/10.3758/BF03204922
Franceschini, S., Gori, S., Ruffino, M., Pedrolli, K., & Facoetti, A. (2012). A causal link between visual spatial attention and reading acquisition. *Current Biology, 22*(9), 814–819. https://doi.org/10.1016/j.cub.2012.03.013
Groppe, D. M., Urbach, T. P., & Kutas, M. (2011). Mass univariate analysis of event-related brain potentials/fields I: A critical tutorial review. *Psychophysiology, 48*(12), 1711–1725. https://doi.org/10.1111/j.1469-8986.2011.01273.x
Harris, C. R., Pashler, H. E., & Coburn, P. (2004). Moray revisited: High-priority affective stimuli and visual search. *The Quarterly Journal of Experimental Psychology, 57*(1), 1–31. https://doi.org/10.1080/02724980343000107
Inhoff, A. W., & Topolski, R. (1992). Lack of semantic activation from unattended text during passage reading. *Bulletin of the Psychonomic Society, 30*(5), 365–366. https://doi.org/10.3758/BF03334090
Jordan, T. R., & Thomas, S. M. (2002). In search of perceptual influences of sentence context on word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*(1), 34–45. https://doi.org/10.1037/0278-7393.28.1.34
Juola, J. F., Ward, N. J., & McNamara, T. (1982). Visual search and reading of rapid serial presentations of letter strings, words, and text. *Journal of Experimental Psychology: General, 111*(2), 208–227. https://doi.org/10.1037/0096-3445.111.2.208
Karlin, M. B., & Bower, G. H. (1976). Semantic category effects in visual word search. *Perception & Psychophysics, 19*(5), 417–424. https://doi.org/10.3758/BF03199402
Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology, 62*(1), 621–647. https://doi.org/10.1146/annurev.psych.093008.131123
Luka, B. J., & Van Petten, C. (2014). Prospective and retrospective semantic processing: Prediction, time, and relationship strength in event-related potentials. *Brain and Language, 135*, 115–129. https://doi.org/10.1016/j.bandl.2014.06.001
Marslen-Wilson, W., & Tyler, L. K. (1980). The temporal structure of spoken language understanding. *Cognition, 8*(1), 1–71. https://doi.org/10.1016/0010-0277(80)90015-3
Pollatsek, A., Raney, G. E., Lagasse, L., & Rayner, K. (1993). The use of information below fixation in reading and in visual search. *Canadian Journal of Experimental Psychology, 47*(2), 179–200. https://doi.org/10.1037/h0078824
Rayner, K. (2009). Eye movements in Reading: Models and data. *Journal of Eye Movement Research, 2*(5), 1–10.
Rayner, K., Inhoff, A. W., Morrison, R. E., Slowiaczek, M. L., & Bertera, J. H. (1981). Masking of foveal and parafoveal vision during eye fixations in reading. *Journal of Experimental Psychology: Human Perception and Performance, 7*(1), 167–179. https://doi.org/10.1037/0096-1523.7.1.167
Rayner, K., Liversedge, S. P., White, S. J., & Vergilino-Perez, D. (2003). Reading disappearing text: Cognitive control of eye movements. *Psychological Science, 14*(4), 385–388. https://doi.org/10.1111/1467-9280.24483
Reichle, E. D., Rayner, K., & Pollatsek, A. (2003). The E-Z reader model of eye-movement control in reading: Comparisons to other models. *Behavioral and Brain Sciences, 26*(4), 445–476. https://doi.org/10.1017/S0140525X03000104
Reichle, E. D., Vanyukov, P. M., Laurent, P. A., & Warren, T. (2008). Serial or parallel? Using depth-of-processing to examine attention allocation during reading. *Vision Research, 48*(17), 1831–1836. https://doi.org/10.1016/j.visres.2008.05.007
Scharff, A., Palmer, J., & Moore, C. M. (2011). Extending the simultaneous-sequential paradigm to measure perceptual capacity for features and words. *Journal of Experimental Psychology: Human Perception and Performance, 37*(3), 813–833. https://doi.org/10.1037/a0021440
Shaffer, W. O., & LaBerge, D. (1979). Automatic semantic processing of unattended words. *Journal of Verbal Learning and Verbal Behavior, 18*(4), 413–426. https://doi.org/10.1016/S0022-5371(79)90228-7
Shaw, M. L. (1978). A capacity allocation model for reaction time. *Journal of Experimental Psychology: Human Perception and Performance, 4*(4), 586–598. https://doi.org/10.1037/0096-1523.4.4.586
Simon, J. R. (1990). The effects of an irrelevant directional CUE on human information processing. In R. W. Proctor, & T. G. Reeve (Eds.), *Vol. 65. Advances in psychology* (pp. 31–86). North-Holland. https://doi.org/10.1016/S0166-4115(08)61218-2.
Snell, J. (2018). OB1-reader: A model of word recognition and eye movements in text reading. *Psychological Review, 1256*, 969–984.
Snell, J., & Grainger, J. (2017). The sentence superiority effect revisited. *Cognition, 168*, 217–221. https://doi.org/10.1016/j.cognition.2017.07.003
Snell, J., & Grainger, J. (2019). Readers are parallel processors. *Trends in Cognitive Sciences, 23*(7), 537–546. https://doi.org/10.1016/j.tics.2019.04.006
Snell, J., Meeter, M., & Grainger, J. (2017). Evidence for simultaneous syntactic processing of multiple words during reading. *PLoS One, 12*(3). https://doi.org/10.1371/journal.pone.0173720
Sperling, G., & Melchner, M. J. (1978). The attention operating characteristic: Examples from visual search. *Science, 202*(4365), 315–318. https://doi.org/10.1126/science.694536
Treisman, A. (1986). Properties, parts, and objects. In handbook of perception and human performance. In *, Vol. 2. Cognitive processes and performance* (pp. 1–70). John Wiley & Sons.
Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12*(1), 97–136. https://doi.org/10.1016/0010-0285(80)90005-5
Underwood, G. (1976). Semantic interference from unattended printed words. *British Journal of Psychology, 67*(3), 327–338. https://doi.org/10.1111/j.2044-8295.1976.tb01519.x
Underwood, G. (1981). Lexical recognition of embedded unattended words: Some implications for reading processes. *Acta Psychologica, 47*(3), 267–283. https://doi.org/10.1016/0001-6918(81)90012-3
Underwood, G., & Thwaites, S. (1982). Automatic phonological coding of unattended printed words. *Memory & Cognition, 10*(5), 434–442. https://doi.org/10.3758/BF03197645
Valdois, S., Bosse, M.-L., & Tainturier, M.-J. (2004). The cognitive deficits responsible for developmental dyslexia: Review of evidence for a selective visual attentional disorder. *Dyslexia, 10*(4), 339–363. https://doi.org/10.1002/dys.284
Vidyasagar, T. R., & Pammer, K. (2010). Dyslexia: A deficit in visuo-spatial attention, not in phonological processing. *Trends in Cognitive Sciences, 14*(2), 57–63. https://doi.org/10.1016/j.tics.2009.12.003
Wen, Y., Snell, J., & Grainger, J. (2019). Parallel, cascaded, interactive processing of words during sentence reading. *Cognition, 189*, 221–226. https://doi.org/10.1016/j.cognition.2019.04.013
White, A. L., Palmer, J., & Boynton, G. M. (2018). Evidence of serial processing in visual word recognition. *Psychological Science, 29*(7), 1062–1071. https://doi.org/10.1177/0956797617751898
White, A. L., Palmer, J., & Boynton, G. M. (2020). Visual word recognition: Evidence for a serial bottleneck in lexical access. *Attention, Perception, & Psychophysics, 82*(4), 2000–2017. https://doi.org/10.3758/s13414-019-01916-z
White, A. L., Palmer, J., Boynton, G. M., & Yeatman, J. D. (2019). Parallel spatial channels converge at a bottleneck in anterior word-selective cortex. *Proceedings of the National Academy of Sciences, 116*(20), 10087–10096. https://doi.org/10.1073/pnas.1822137116
White, M. J. (1969). Laterality differences in perception: A review. *Psychological Bulletin, 72*(6), 387–405. https://doi.org/10.1037/h0028343
Williams, P. C., & Parkin, A. J. (1980). On knowing the meaning of words we are unable to report confirmation of a guessing explanation. *Quarterly Journal of Experimental Psychology, 32*(1), 101–107. https://doi.org/10.1080/00335558008248236
Wolfe, J. M. (1994). Guided search 2.0 a revised model of visual search. *Psychonomic Bulletin & Review, 1*(2), 202–238. https://doi.org/10.3758/BF03200774
Yantis, S., & Johnston, J. C. (1990). On the locus of visual selection: Evidence from focused attention tasks. *Journal of Experimental Psychology: Human Perception and Performance, 16*(1), 135–149. https://doi.org/10.1037/0096-1523.16.1.135