Conflict resolution ability underpins the processing of temporary syntactic ambiguity: Evidence from ERPs.

Polly O'Rourke (porourke@casl.umd.edu)
Center for Advanced Study of Language
University of Maryland, 7005 52nd Avenue
College Park, MD 20742 USA

Gregory Colflesh (gcolflesh@casl.umd.edu)
Center for Advanced Study of Language
University of Maryland, 7005 52nd Avenue
College Park, MD 20742 USA

Abstract

The resolution of temporary syntactic ambiguity involves not only the reanalysis of syntactic structure but also the resolution of the conflict between the new correct interpretation and the initial incorrect parse. The current study investigated the contribution of domain-general conflict resolution ability as indexed by performance on an *n*-back task (with lures) and the online resolution of temporary syntactic ambiguity as indexed by the P600. P600s for temporarily ambiguous sentences were predicted by accuracy for $n + 1$ lures. No such relationship was found for unambiguous, syntactically complex sentences. These results suggest that domain-general conflict resolution ability underpins the online resolution of temporary syntactic ambiguity.

Keywords: Garden-Path; Conflict Resolution, P600, *N*-back, Working Memory.

Introduction

Sentences containing temporary syntactic ambiguities (i.e., garden-path sentences) represent a unique challenge during parsing as they typically lead to the reanalysis of an initial preferred but incorrect parse. For example in sentence (1) below *the nurse* is typically initially interpreted as conjoined with the noun phrase *the patient* and thus the direct object of the verb *met*. Upon encountering the verb *showed* it is obvious that this interpretation is incorrect and the structure must be revised such that *the nurse* is the subject of a conjoined sentence.

(1) The doctor met the patient and the nurse with the white coat showed the chart at the meeting.

This reanalysis process is costly and not always successful, thus resulting in longer reading times and reduced comprehension accuracy (Frazier & Rayner, 1982). Research using the event-related potential (ERP) technique has also indicated that the resolution of temporary syntactic ambiguity places increased demands on the parser such that the disambiguating word in a garden-path sentence (*showed* above) elicits the P600 effect. The P600 (a positivity maximal between 500 and 800 ms post stimulus over centro-parietal areas) is an ERP generally associated with difficulty in syntactic processing such that it is elicited by syntactic violations of all types (Kutas, van Petten & Kluender, 2006) but also by well-formed sentences that present increased difficulty due to temporary ambiguity (Gouvea, Phillips, Kazanina & Poeppel, 2010; Osterhout, Holcomb & Swinney, 1994) or syntactic complexity (Kaan, Harris, Gibson & Holcomb, 2000).

One explanation for the increased difficulty associated with resolving temporary syntactic ambiguity is that the structural reanalysis places an increased demand on cognitive control abilities (i.e., working memory). Working memory (WM) is "a multicomponent system responsible for active maintenance of information in the face of ongoing processing and/or distraction" (Conway, Kane, Bunting, Hambrick, Wilhelm & Engle, 2005, p. 770) that facilitates goal directed behavior. Individual differences in working memory capacity (WMC) impact sentence processing ability (Just & Carpenter 1992). Indeed, many studies have shown that working memory capacity (WMC) predicts garden-path effects in both online and offline measures. Individuals with high WMC show greater comprehension accuracy for garden-path sentences compared to their low WMC counterparts (Just & Carpenter, 1992). Furthermore, individuals with greater WMC show greater P600 effects for garden-path sentences (Friederici et al., 1998; O'Rourke, 2013; Vos & Friederici, 2003; Bornkessel et al., 2004; Vos, Gunter, Schriefers & Friederici et al., 2001) suggesting that their increased capacity helps them with the structural reanalysis.

Understanding of the relationship between WMC and sentence processing is limited by the fact that the overwhelming majority of studies use one type of WM assessment. Language processing studies typically use a complex span task, usually the reading span task (Danemann & Carpenter, 1980; Unsworth, Heitz, Schrock & Engle, 2005) as the sole index of WMC. In the reading span task, participants are presented with a sequence of items (typically words or letters) for subsequent serial recall. After the presentation of each to-be-remembered item in the sequence, participants are presented with a sentence about which they must make a judgment (e.g., a sense or grammaticality judgment). The reading span task assesses an individual's ability to store and recall information during the processing of additional interfering information (Conway, et al., 2005). In this way, the reading span task is akin to garden-path resolution as reanalyzing a garden-path structure requires storage of the linear sequence of words with concurrent recreation of syntactic structure (i.e., processing), hence a possible explanation for the observed relationship between complex span tasks and garden-path resolution. However, given that the processing component of the reading span task (i.e., reading sentences) is very closely, if not directly, related to reading comprehension ability it is difficult to determine if it is WMC or reading ability that mediates the relationship between reading span and sentence processing.

There are, however, a multitude of WM measures which tap into to other executive functions (EFs) (Conway et al., 2005). One task predominately used in cognitive psychology, cognitive neuroscience, and cognitive training research is the *n*-back task (Conway et al., 2005; Kane, Conway, Miura & Colflesh, 2007). *N*-back performance reflects the ability to maintain, monitor and regularly update information, and in some versions of the task, to resolve the conflict of competing representations. In the *n*-back task, participants monitor a sequence of stimuli, typically letters, and indicate whether the current stimulus matches the one the appeared *n* spaces back in the sequence. Often lure items are included which match the target stimulus, but do not appear in the target position (e.g. matching the letter that appears 3 spaces back in a 4-back task). The inclusion of lures creates conflict between familiar and recollected information. Novick, Hussey, Teubner-Rhodes, Dougherty, Harbison & Bunting (2013) found that individuals who improve *n*-back (with lures) performance via training show reduced garden-path effects in terms of both reading times and comprehension accuracy. They concluded that resolving conflict between competing mental representations is the central domain-general mechanism to both lure performance and garden-path resolution. O'Rourke (2013) found that *n*-back performance predicted effects of syntactic complexity such that the P600 and accuracy effects for unambiguous object relatives correlated with *n*-back lure accuracy while garden-path effects correlated with complex span performance. O'Rourke concluded that given that the resolution of the long distance filler-gap dependency of the object relative requires monitoring each incoming word to determine if it can be matched to the extracted wh-word, resolving filler-gap dependencies is similar in task demands to the *n*-back task. Though the divergent findings are perplexing, it is difficult to directly compare these two studies given the extensive methodological differences.

The goal of the current study was to further examine the relationship between conflict resolution and temporary syntactic ambiguity by examining how different WM measures predict ERPs and comprehension accuracy for garden-paths and object relatives. By addressing certain methodological concerns from O'Rourke (2013) (see methods), the current study aimed to elucidate the relationship between conflict resolution and garden-path reanalysis in both online (P600) and offline (comprehension accuracy) processing. While sentence type effects were assessed with standard repeated measures analysis of variance (ANOVA), the relationship between EF and measures of sentence processing were assessed with multiple regression.

Methods

Participants

Data were collected from 62 right handed participants, 25 of which were excluded due to eligibility issues, technical issues, noncompliance or excessive artifacts. As a result, 37 participants (22 female) between the ages of 18 and 35 were included in the analysis. All participants were right-handed, neurologically normal, native speakers of English with normal or corrected-to-normal vision, and none had had started learning a second language before age 12.

Sentence Stimuli

For the syntactic complexity conditions, this experiment used the same sentence types as O'Rourke (2013) (based on Gouvea, et al., 2010; see Table 1). Fifty percent of the questions were followed by comprehension questions. The comprehension questions were modified to include questions directly targeting the resolution of the garden-path structure. The words to which the ERP data is time-locked are underlined in the examples below. There were 36 sentences per condition in the each syntactic complexity conditions making a total of 108 target sentences.

Table 1. Examples of Sentence Types

Type	**Example**
Garden-Path	The patient met the doctor and the nurse with the white dress <u>showed</u> the chart during the meeting.
Object Relative	The patient met the doctor to whom the nurse with the white dress <u>showed</u> the chart during the meeting.
Control	The patient met the doctor while the nurse with the white dress <u>showed</u> the chart during the meeting.
Question	Did the patient meet the nurse?

O'Rourke (2013) included only 72 filler sentences and did not systematically include distractor sentences to reduce the predictability of the target structures. In the current experiment, categories of fillers were included for this purpose (see Table 2). These strategic filler conditions included sentences with conjoined noun phrases in the direct object position (as distractors for the garden-path sentences), object relatives in which the extracted wh-phrase ("who") was the direct object of the verb in the relative clause, and indirect object relatives with "who" as the relative pronoun (as distractors for the object relatives). There were 36 sentences in each of these filler conditions resulting in 108 strategic filler sentences. An additional 144 filler sentences were added (including two experimental conditions not reported herein) making a total of 360 sentences. Due to the large number of sentences, the sentence processing task was divided into two sessions.

Table 2. Examples of Filler Sentences

Sentence Type	**Example**
Object Relative – Direct Object	The man washed the dog who the child with the muddy clothes brought in the house.
Object Relative – Indirect object with "who"	The manager avoided the woman who the dealer with the horrible tie sold the car to in the garage.
Conjoined Noun Phrase	The teacher told the girl and the boy with the failing grade to get extra help.

Working Memory Tasks

Complex Span Tasks: Automated Operation Span (Unsworth, et al., 2005), Automated Reading Span (Unsworth, Redick, Heitz, Broadway & Engle, 2009) and Automated Symmetry Span (Unsworth, et al, 2009) were used in this experiment (see also Redick, Broadway, Meier, Kuriakose, Unsworth, Kane & Engle 2012). All three complex span tasks required participants to maintain memoranda in the focus of attention while performing an irrelevant processing task (e.g., verifying the accuracy of math problems, judging whether a sentence makes sense, and whether pictures are vertically symmetrical for operation, reading, and symmetry spans, respectively).

***N*-Back** The *n*-back task with lures (Conway et al., 2005; Kane, Conway, Miura & Colflesh, 2007) was used as an index of conflict resolution ability. Participants performed 2-back and 4-back. Accuracy for $n + 1$ and $n - 1$ lures in the 4-back task were used as predictor variables in the regression analysis. The selection was based on the findings of Kane et al. and Novick et al. (2013) that suggest that accuracy on 4-back lures are the most likely to be dependent on conflict resolution ability.

Simon Task The Simon task (Simon, 1990) measures the index of the ability to suppress a prepotent, but goal-irrelevant response. Performance was measured by the difference in reaction times between accurate incongruent and congruent trials.

EEG Recording

Electroencephalographic (EEG) data was acquired using the Electrical Geodesics Inc. (EGI) Hydrocel 256 channel system. Data were recorded using Net Station 4.5.4 (Electrical Geodesics Inc., Eugene, OR). The signal was high-pass filtered online at 0.1 Hz, low-pass filtered at 30 Hz. The EEG signal was sampled at 250 Hz. Impedances were kept below 50 kΩ where possible and otherwise under 100 kΩ. EEG was recorded using CZ as a reference and later re-referenced to the global mean. Prior to averaging, drift, eye blinks and other movement artifacts were identified and excluded from analysis via visual inspection of the data. Participants with more than 40% rejected trials were not included in the analysis. In the subjects included in the analysis, 20% of trials were rejected on average.

Procedure

The sentence processing task was broken into two sessions, each lasting approximately 2.5 hours. In each session, participants performed the sentence processing task while EEG was recorded, and then they completed a portion, either two or three, of the five WM assessments. In the second session, they completed the remainder of the five WM assessments. Sentences appeared word-by-word in the center of a high-resolution computer screen. Each word was presented for 300 ms, followed by a blank of 200 ms. The final word of the sentence was presented with a period sign and was followed by a 5.5 second rest period. Half of the test sentences were followed by a comprehension question. The questions were presented in their entirety for 2500 ms, followed by a rest period of 3500 ms. Key presses were used for yes and no responses to the questions. Within each session, the stimuli were broken into 6 runs consisting of 27 sentences, lasting approximately 8 minutes each. The EEG session, including electrode application and removal, lasted approximately 1.5 hours. After electrode removal, participants performed two or three of the five working memory assessments (also in a sound-attenuated booth) which took no longer than one hour.

Data Analysis

Data were analyzed using Net Station 4.5.4 (Electrical Geodesics Inc., Eugene, OR). Upon completion of pre-processing and averaging, ERPs were computed for each individual in each experimental condition for a 1500 ms interval time-locked to the presentation of the critical verb ("showed" above) relative to a 200 ms pre-stimulus baseline. The following time windows were considered in the analysis: 500-700 and 700-900 ms. The analyses were performed on midline, dorsal and ventral electrodes (Coulson & Van Petten, 2007; O'Rourke & Van Petten, 2011; Thornhill & Van Petten, 2012). The midline electrodes were divided into anterior (FPZ, AFZ, FZ, FCZ, CZ) and posterior (CPZ, 90, PZ, POZ, OZ) sections. The dorsal electrodes were grouped by anterior-posterior (AP) location and hemisphere: Left anterior (FP1, AF3, F1, F3, FC3, C3), right anterior (FP2, AF4, F2, F4, FC4, C3), left posterior (CP3, CP1, P1, P3, P1, PO3, O1) and right posterior (CP4, CP1, P4, P2, PO4, O2). The ventral electrodes were similarly grouped: Left anterior (F7, F5, FT7, FC5, T3, C5), right anterior (F8, F6, FT8, FC6, T4, C6), left posterior (TP7, CP5, P5, T5, PO7, P9) and right posterior (TP8, CP6, P6, T6, PO8, P10).

Sentence type effects in the ERP data were assessed in the dorsal and ventral regions with three-way ANOVAs (sentence type x AP x hemisphere) and in the midline electrodes with a two-way ANOVA (sentence type x AP).

Four stepwise linear regression analyses were run using the mean amplitude in posterior midline electrodes in the 700-900 ms time window and comprehension accuracy for garden-path and object relative sentences, and comprehension accuracy for both sentence types as the dependent variables. In addition to the accuracies for the two 4-back lure conditions referenced above and the Simon congruency effect on RTs, a composite complex span score (i.e., average across the three tasks) was included as a predictor variable in the regression analysis. Additional regression analyses were run for object relative sentences using the mean amplitude over anterior and posterior right ventral areas (700-900 ms). For all behavioral data, any data points ± 2.5 standard deviations (SD) from the sample mean were excluded from the analysis. As a result data from 33 participants were included in the regression analyses.

Results

Behavioral Data

Accuracy for Garden-Path (M = 68.6%, SD = 14.2) and object relative sentences (M = 66.8%, SD = 12.1) were lower than control sentences (M = 74.1%, SD = 10.8). Both sentence type effects were significant: Garden-path versus controls ($F(1,36) = 6.32$, $p < .05$) and Object relatives versus controls ($F(1,36) = 9.89$, $p < .005$). Average complex span scores were as follows: Operation span (M = 53.9, SD = *14*.0), reading span (M = 40.9, SD = 15.7) and symmetry span (M = 20.9, SD = 15.7). Mean RTs for congruent and incongruent Simon trials were 411 ms (SD = 46) and 439 (SD = 46), respectively. Mean accuracies for n + 1 and n - 1 lures were 67.9% (SD = 24.5) and 67.1% (SD = 28.2), respectively.

ERP Data

Garden-path sentences elicited greater positivity than controls over posterior areas. This effect emerged in the 500-700 ms time window as an interaction of sentence type by AP over midline sites ($F(1,36) = 4.75$, $p < .05$). Splitting across the factor AP revealed a simple effect of sentence type over posterior sites only ($F(1,36) = 4.57$, $p < .05$). Similarly, in the 700-900 ms window, there was a significant interaction of sentence type by AP over midline sites ($F(1,36) = 4.37$, $p < .05$) driven by a simple effect of sentence type in posterior sites ($F(1,36) = 4.32$, $p < .05$). Object relatives elicited increased positivity compared to controls but the effect was limited to the 700-900 ms time window over ventral areas. The effect manifested as an interaction of sentence type and hemisphere ($F(1,36) = 4.83$, $p < .05$), driven by a simple effect of sentence type in the right hemisphere ($F(1,36) = 4.52$, $p < .05$) such that object relatives elicited greater positivity over the right hemisphere.

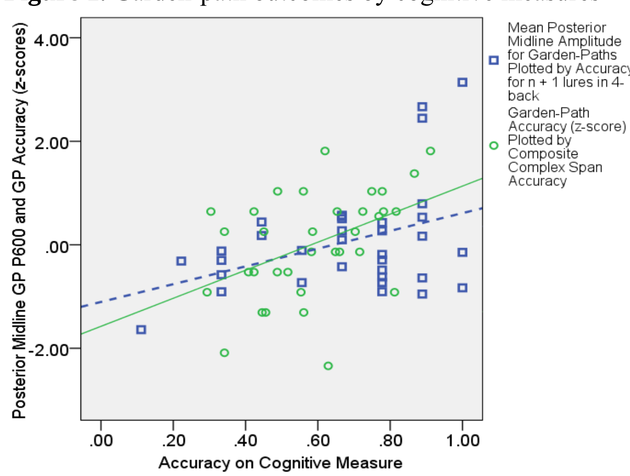
Correlations among WM Measures

There was a significant correlation between symmetry span and reading span ($r = .50$, $p < .05$). There was a marginal correlation between the Simon congruency effect and accuracy for n - 1 lures ($r = -.31$, $p = .06$) such that better lure accuracy was associated with smaller differences between incongruent and congruent trials. Critically, there were no correlations between the complex span tasks and n-back performance which is consistent with previous findings (Kane, et al., 2007; Unsworth, Schrock & Engle, 2004).

Regression Analyses

A stepwise multiple regression analysis was performed using the mean amplitude over posterior midline sites in the 700-900 ms time window for garden-path sentences as the dependent variable with n + 1 lure accuracy, n - 1 lure accuracy (both in the 4-back task), mean complex span score, and Simon congruency effect as predictor variables. N + 1 lure accuracy was the only significant predictor of P600 amplitude for garden-path sentences (adjusted R^2 = .13, $F(1,32) = 6.07$, $p < .05$; $\beta = .40$, $p < .05$). Comprehension accuracy for garden-paths was predicted by Mean Composite Complex Span score (adjusted R^2 = .14, $F(1,32) = 6.48$, $p < .05$; $\beta = .41$, $p < .05$). Scatterplots illustrating these two outcomes are presented in Figure 1.

Figure 1. Garden-path outcomes by cognitive measures



Note: The dashed line represents the ERP regression fit line. The solid line represents the behavioral fit line.

None of the other independent variables were significant predictors in this model. Analyses conducted for object relatives revealed no significant predictors for late posterior positivity or comprehension accuracy.

Discussion

The effects of sentence type in both the ERPs and comprehension accuracy are consistent with previous findings (Frazier & Rayner, 1982; Frazier, 1987; King & Just, 1991). In O'Rourke (2013) the effect for garden-paths was marginal (p = .05) and not significant for object relatives (p > .2). The methodological adjustments to reduce predictability do, therefore, seem to have accentuated the sentence type effects. Likewise, the P600 for garden-paths has an earlier latency and longer duration than in O'Rourke (2013) and are more similar to Gouvea et al. (2010). In the current results there are clear topographical distinctions between the late positivities associated with resolving temporary syntactic ambiguity and filler-gap dependencies. Though not the focus of this paper, this finding supports O'Rourke (2013)'s conclusion that the late positive components for the two sentence types are functionally distinct and the characterization of the P600 as a collection of late positivities reflecting distinct mechanisms in sentence processing.

The results of the regression analysis confirm the connection proposed by Novick et al. (2013) between domain-general conflict resolution ability and the online reanalysis of garden-path sentences such that individuals who performed better on n + 1 lures in the 4-back task showed increased posterior midline positivity during the 700-900 ms time window for garden-path sentences. Friederici et al. (1998)'s account of the relationship between WMC and garden-path resolution (based on reading span performance) have emphasized the difference in parsing efficiency such that "[h]igh span readers are more efficient parsers than low span readers because they commit themselves to a single preferred structure when confronted with structural ambiguities" (Friederici et al., 1998, p. 219). Encountering the disambiguating word compels the high span individuals to undergo the costly reanalysis process while perhaps low spans keep all possible parses active. The current results do not contradict this account but indicate that the critical factor determining the character of the online reanalysis is the ability to manage the conflict between the two active representations, only one of which is correct.

This finding is, however, a clear deviation from previous findings of O'Rourke (2013)'s correlational study as well as Friederici et al. (1998), Vos et al. (2001), and Bornkessel et al. (2004), all of which found a relationship between P600 effect size and complex span performance. With respect to O'Rourke (2013), one key difference is that in O'Rourke (2013) garden-path sentences constituted 20% of the sentence stimuli and in the current study they were only 10%. It is known that the probability of a sentence type in the stimuli set modulates the P600 such that low probability sentences elicit a P600 effect regardless of grammaticality (Coulson, King and Kutas, 1998). In both Friederici et al. (1998) and Vos et al. (2001), both of which found that reading span performance related to the P600 for garden-paths, the probability of temporary syntactic ambiguity was 50%. Bornkessel et al. (2004) got findings similar to Friederici et al. (1998) and Vos et al. (2001) findings with 25% temporary syntactic ambiguity. In addition to the lower percentage, the current study also included deliberate distractors containing noun phrase conjunction to reduce the predictability of the garden-path structure. This may suggest that when temporarily ambiguous sentences are more predictable, reanalysis involves less conflict resolution and increased attention to the linear sequence during processing. When probability is very low, garden-path effects are stronger and, thus, online reanalysis is predicted by conflict resolution ability.

The lack of a relationship between conflict resolution ability and the late positivity for object relatives (neither in the posterior midline, nor the right ventral areas) is not surprising as these sentences require the completion of a filler-gap dependency in absence of ambiguity. While further research is required to map out the connection with working memory and the processing of complex syntax (especially as these results contradict those of O'Rourke, 2013), this finding suggests that the difficulty associated with processing object relatives is not due to increased working memory demand.

Given that, as in previous studies (Just & Carpenter, 1992; King & Just, 1991) complex span performance predicts comprehension accuracy for garden-path sentences, the current results suggest a clear distinction between the cognitive demands of online and offline processing such that online processing of garden-paths reflects the process of reanalysis (and resolving conflict among competing representations) while offline may reflect the ability to recall the correct representation (rather than construct/select it).

In conclusion, the results of the current study provide evidence that the reanalysis of syntactic structure associated with resolving a garden-path is underpinned by a domain general conflict resolution mechanism. Furthermore, the demands of resolving garden-path structures are associated with neurocognitive mechanisms distinct from those associated with resolving filler-gap dependencies. Our results also suggest to a divide between on-line and off-line processing in terms of cognitive demand.

References

Bornkessel, I.D., Fiebach, C.J. & Friederici, A.D. (2004). On the cost of syntactic ambiguity in human language comprehension: an individual differences approach. *Cognitive Brain Research*, 21, 11-21.

Conway, A.R.A., Kane, M.J., Bunting, M.F., Hambrick, D.Z., Wilhelm, O. & Engle, R.W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review*, 12(5), 769-786.

Coulson, S., King, J.W. & Kutas, M. (1998). Expect the unexpected: Event-related brain responses to morphosyntactic vioations. Language and Cognitive Processes, 13(1), 21-58.

Daneman, M., & Carpenter, P.A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, 19, 450–466.

Frazier, L. (1987). Syntactic processing: Evidence from Dutch. *Natural Language and Linguistic Theory*, 5, 519–559.

Frazier, L. & Rayner, K. (1982). Making and correcting errors during sentence comprehension: eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology*, 14, 178-210.

Friederici, A.D., Steinhauer, K., Mecklinger, A., & Meyer, M. (1998). Working memory constraints on syntactic ambiguity resolution as revealed by electrical brain responses. *Biological Psychology*, 47, 193–221.

Gouvea, A.C., Phillips, C., Kazanina, N. & Poeppel, D. (2010). The linguistic processes underlying the P600. *Language and Cognitive Processes*, 25(2), 149-188.

Just, M.A., & Carpenter, P.A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, 99, 122–149.

Kaan, E., Harris, T., Gibson, E., Holcomb, P.J., 2000. The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes*. 15, 159–201.

Kane, M.J., Conway, A.R.A., Miura, T.K. & Colflesh, G.J.H. (2007). Working memory, attention control, and the N-back task: A question of construct validity. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 33(3), 615-622.

King, J.W., & Just, M.A. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language*, 30, 580–602.

Kutas, M., Van Petten, C., & Kluender, R. (2006). Psycholinguistics electrified II: 1994-2005. In M. Traxler & M.A. Gernsbacher (Eds.), *Handbook of Psycholinguistics, 2nd Edition*. New York: Elsevier.

Novick, J.M., Hussey, E., Teubner-Rhodes, S., Dougherty, M.R., Harbison, J.I., & Bunting, M.F. (2013). Clearing the garden-path: Improving sentence processing through executive control training. *Language and Cognitive Processes,* (ahead of print), 1-44.

Novick, J.M., Trueswell, J.C. & Thompson-Schill, S.L. (2005). Cognitive control and parsing: Reexamining the role of Broca's area in sentence comprehension. *Cognitive, Affective & Behavioral Neuroscience*, 5(3), 263-281.

O'Rourke, P. (2013). The interaction of different working memory mechanisms and sentence processing: A study of the P600. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the 35th Annual Conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.

Osterhout, L., Holcomb, P.J., Swinney, D.A., 1994. Brain potentials elicited by garden-path sentences: evidence of the application of verb information during parsing. *Journal of Experimental Psychology: Learning Memory and Cognition*, 20, 786–803.

Redick, T.S., Broadway, J.M., Meier, M.E., Kuriakose, P.S., Unsworth, N., Kane, M.J., & Engle, R.W. (2012). Measuring working memory capacity with automated complex span tasks. *European Journal of Psychological Assessment*, *28*, 164-171.

Simon, J.R. (1990). The effects of an irrelevant directional cue on human information processing. In R. W. Proctor & T. G. Reeve (Eds.), *Stimulus-response compatibility: An integrated perspective* (pp. 31–86). Amsterdam: North-Holland.

Sprouse, J., Wagers, M. & Philips, C. (2012). A test of the relation between working-memory capacity and syntactic island effects. *Language*, 88(1), 82-123.

Turner, M.L. & Engle, R.W. (1989). Is working memory capacity task independent? *Journal of Memory and Language*, 28, 127-154.

Unsworth, N., Schrock, J.C. & Engle, R.W. (2004). Working memory capacity and the antisaccade task: Individual differences in voluntary saccade control. *Journal of Experimental Psychology: Learning Memory and Cognition*, 30, 1302-1321.

Unsworth, N., Heitz, R.P., Schrock, J.C., & Engle, R.W. (2005). An automated version of the operation span task. *Behavior Research Methods, 37*, 498-505.

Vos, S.H., Gunter, T., Schriefers, H. y Friederici, A.D. (2001). Syntactic parsing and working memory: The effects of syntactic complexity, reading span and concurrent load. *Language and Cognitive Processes*, 16, 65-103.

Acknowledgements

This research is based upon work supported, in whole or in part, with funding from the United States Government. Any opinions, findings and conclusions, or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the University of Maryland, College Park and/or any agency or entity of the United States Government. Nothing in this article is intended to be and shall not be treated or construed as an endorsement or recommendation by the University of Maryland, United States Government, or the author of the product, process, or service that is the subject of this article. Correspondence concerning this article should be addressed to Polly O'Rourke, Center for Advanced Study of Language, University of Maryland, 7005 52nd Avenue, College Park, MD 20742 (e-mail: porourke@casl.umd.edu). Special thanks to Mike Bunting and Valerie Karuzis.