Response Dominance Predicts Garden-Path Comprehension: An ERP Study

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

Abstract

While the P600 is generally presumed to be a uniform response elicited consistently across individuals in specific syntactic contexts, Tanner and Van Hell (2014) showed evidence of distinct response profiles (N400 or P600 dominant) for syntactic violations across individuals. The current analysis used Tanner and Van Hell's response-dominance index (RDI) to examine the impact of response dominance on comprehension of garden-path sentences. P600 dominant individuals showed enhanced comprehension of garden-path sentences, even when controlling for working memory capacity. Response dominance as an individual difference measure has the potential to enhance understanding of the neurocognitive basis of sentence processing and greater cognition in general.

Keywords: P600; N400; Response Dominance Index; Garden-Path; Working Memory.

Introduction

Over the past thirty years, research using the event-related potential (ERP) technique has provided many important insights into the neural mechanisms associated with language comprehension. The integration of lexico-semantic information is associated with an increased centro-parietal negativity between 300-500 ms known as the N400 (Kutas & Hillyard 1980). Morphosyntactic integration is associated with an early left anterior negativity (LAN) maximal around 200-500 ms, followed by a late posterior positivity (P600) maximal between 500 and 800 ms (see Kutas, Van Petten & Kluender (2005) for review). While the LAN/P600 complex is elicited by morphosyntactic violations of all types, the P600, in absence of early negativity, is also elicited by well-formed sentences that present increased difficulty due to temporary ambiguity (i.e. garden-paths; Gouvea, Phillips, Kazanina & Poeppel, 2010; Osterhout, Holcomb & Swinney, 1994).

The use of ERP as a means of indexing the different neural mechanisms associated with language processing is contingent on the assumption that all neurologically normal, native speakers show consistent responses to sentence stimuli such that the grand averaged ERPs reflect effects that are manifest uniformly across individuals.

This notion was recently challenged by Tanner and Van Hell (2014). In their innovative study, they showed that, although in the grand mean syntactic violations elicited a classic biphasic LAN/P600 response, most participants either showed an N400 or a P600 rather than a biphasic response. Given the topographical distribution of the effects for each group, they concluded that the LAN often found for syntactic violations in grand mean analyses is the result of the distributed negativity in some subjects being neutralized or minimized by the right lateralized positivity in the others such that only the left anterior negativity remains. Response dominance did not, however, predict acceptability judgment accuracy, nor did it correlate with measures of working memory (WM) and executive control. Familial left-handedness alone exhibited a relationship with the response dominance index (RDI).

The individual differences in N400/P600 response dominance observed by Tanner and Van Hell (2014) led to interesting questions regarding other contexts which tend to elicit these potentials. Although sentences that are well-formed but contain garden-paths typically elicit a P600 with the absence of an early negativity (Friederici, Mecklinger, Spencer, Steinhauer, & Donchin, 2001; Frisch, Schlesewsky, Saddy, & Alperman, 2002; Gouvea et al., 2010; Horberg, Koptjevskaja, & Kallioinen, 2013; Matzke, Mai, Nager, Rossler, & Münte, 2002; Osterhout, Holcomb & Swinney, 1994; Vos et al., 2001a), there is some variability as many studies do observe early effects. Horberg et al. (2013) found early centro-parietal negativity while Matzke et al. (2002) found early anterior negativities. Early positivities have also been found, both posterior (Gouvea et al., 2010; Vos et al., 2001a) and broadly distributed (Friederici et al., 2001)

Though the effects associated with the resolution of garden-paths are not considered to be biphasic, the variability in findings could suggest the possibility of individual differences in response profiles, as Tanner and Van Hell (2014) found for syntactic violations. Furthermore, even if all studies found P600s in absence of negativity, the lack of early effects could be the result of early negativities and positivities cancelling each other out.

One known source of variability in garden-path effects for both P600s and comprehension accuracy is working memory capacity (WMC). 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) which facilitates goal directed behavior. High WMC individuals show greater P600 effects for garden-path sentences compared to low WMC individuals (Friederici, Steinhauer, Mecklinger, & Meyer, 1998). High WMC individuals also show reduced garden-path effects in comprehension accuracy such that they have better comprehension accuracy for garden-paths (Just & Carpenter, 1992). Lower comprehension accuracy in low WMC individuals indicates that they are more likely to arrive at "Good Enough" interpretations (Ferreira, Bailey & Ferraro, 2002) in which the faithful interpretation of the sentence is not adopted. As Slattery, Sturt, Christianson, Yoshida and Ferreira (2013) note, this could be due to a failure to fully reanalyze the structure created during the initial incorrect parse or to a failure to discard the initial incorrect interpretation. Both possibilities represent distinct attentional demands, specifically the ability to recall information in the face of ongoing processing (reanalysis) and the ability to discard/suppress information that is no longer goal relevant (discarding incorrect parse). The former can be assessed with complex span tasks (Daneman & Carpenter, 1980; Unsworth, Heitz, Schrock, and Engle, 2005) while the latter can be indexed by performance on positive lures in the *n*-back task (Conway et al., 2005). It is necessary, therefore, to consider the role of WM when examining individual differences associated with the resolution of garden-path sentences.

The goal of the current study is to determine (1) if participants' N400/P600 dominance for garden-path sentences will fall into a continuum such that there will be a continuous distribution of N400 and P600 effect magnitudes with negative correlations between them, (2) if the N400 and P600 dominant participants will show qualitatively different sentence type effects in the ERP data, (3) if response dominance will predict comprehension accuracy and (4), if so, is that effect reducible to individual differences in WMC.

Methods

Participants

Data were collected from 60 right handed participants, 15 of which were excluded due to eligibility issues, technical issues, noncompliance, or attrition. An additional 8 were excluded for excessive artifacts. As a result, 37 participants (20 female) between the ages of 18 and 35 ($M = 21.6$, $SD = 3.21$) were included in the analyses. All participants were right-handed, neurologically normal, native speakers of English with normal or corrected-to-normal vision, and none had started learning a second language before age 12.

Sentence Stimuli

For the syntactic complexity conditions, this experiment used the same control and garden-path sentences as O'Rourke & Colflesh (2014) (based on Gouvea et al., 2010; see Table 1).

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.
Control	The patient met the doctor while the nurse with the white dress <u>showed</u> the chart during the meeting.

Fifty percent of the sentences were followed by comprehension questions, which included questions that directly targeted the resolution of the garden-path structure. ERP data was time-locked to a matrix verb (underlined in Table 1).

In addition to the target sentences, there were 288 other sentences (108 fillers and 180 sentences in experimental conditions not reported herein; see O'Rourke & Colflesh 2014 for details on fillers) making a total of 360 sentences. Due to the large number of sentences, the sentence processing task was divided across two experimental sessions.

Working Memory Tasks

Complex Span Tasks Three complex span tasks were used in the current study: reading span, operation span, and symmetry span. In automated reading span (Unsworth, et al., 2005) participants were presented with a series of sentences and asked to indicate, via button press, if the sentence they read made sense. After each sentence they were then presented with a letter that they were to remember for later recall. At the end of the sequence, they had to recall the letters in serial order. Their score reflects the total number of letters recalled in the correct serial position out of a total of 75 items.

Automated operation span (Unsworth, Redick, Heitz, Broadway, & Engle, 2009) was identical to reading span as described above except instead of making sense judgments on sentences, participants had to read math problems involving two operations, one addition/subtraction and one multiplication/division, and verify if the solution provided was correct.

In automated symmetry span (Unsworth et al., 2009) participants were presented with a series of 8x8 black and white grids and asked to indicate, via button press, whether the design was vertically symmetrical. After each symmetry judgment they were presented with a 4x4 grid with a square filled in red that they were asked to remember for later recall. At the end of the sequence, participants had to recall the position of the red squares in serial order. Their score reflects the total number of red squares recalled in the correct serial position out of a total of 42.

***N*-Back** In the *n*-back task (Conway et al., 2005), participants were presented with a sequence of single letters and asked to judge if the current letter is the same as the one that occurred *n* places back in the sequence. For example, in a 2-back task, the second "X" in the following sequence would be a target: X U X X F U U. Lures, which appeared one space before a target (n-1; the third "U"), one space after a target (n+1; the third "X"), or two spaces after a target (n+2; the second "U") were also included. Participants in the current experiment performed 2-back and 4-back, with lures. Accuracy in the positive lure conditions reflects the ability to successfully discard recently activated information that is no longer goal relevant. Mean accuracy for n+1 and n+2 lures in the 2-back task was used as a covariate in the ANCOVA.

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, as is customary with high density EEG recordings. EEG was recorded using CZ as a reference and later re-referenced to the global mean. Prior to averaging, eye blinks, and other artifacts were identified and excluded from analysis via Net Station's algorithms for artifact detection and ocular artifact removal. Bad channels were corrected when possible via NetStation's correction tool. Participants with more than 30% rejected trials were not included in the analysis resulting in the exclusion of eight participants. In the subjects included in the analysis, 15% of trials were rejected on average.

Procedure

The sentence processing task was divided across 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 subset of the WM assessments (the three complex span tasks, *n*-back, and two additional measures not reported herein). 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 and was followed by a 5500 ms 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 3500 ms rest period. Key presses were used for "yes" and "no" responses to the questions. Within each session, the stimuli were broken into 6 runs each consisting of 27 sentences, lasting approximately 8 minutes. The EEG part of the session, including electrode application and removal, lasted approximately 1.5 hours. After electrode removal, participants performed three of the six WM assessments 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 for 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 of P600 effects: 300-500, 500-700 and 700-900 ms. The analyses were performed on midline and dorsal electrodes. 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).

Sentence type effects in the ERP data were assessed in the dorsal regions with multiple three-way analysis of variance (ANOVA) (sentence type (Type) x anterior-posterior (AP) x hemisphere) and in the midline electrodes with a two-way ANOVA (Type x AP).

In order to examine the possibility of qualitatively distinct electrophysiological effects across participants, we used Tanner and Van Hell (2014)'s RDI to sort participants into negativity and positivity dominant groups. First, N400 and P600 effect magnitudes for the garden-path/control contrast were calculated using the same parameters as Tanner and Van Hell (p. 293). N400 (control minus garden-path) and P600 effect magnitudes (garden-path minus control) were calculated in the 300-500 and 500-800 ms time windows, respectively, in a centro-parietal ROI (C3, CZ, C4, CP1, CP2, P3 PZ, P4). RDI was calculated for each sentence type using Tanner and Van Hell's formula:

$$\text{RDI} = \frac{(\text{P600}_{\text{GP}} - \text{P600}_{\text{Control}}) - (\text{N400}_{\text{Control}} - \text{N400}_{\text{GPl}})}{\sqrt{2}}$$

To examine the impact of RDI on comprehension accuracy, participants were divided into groups according to response dominance (N400 or P600). An analysis of covariance (ANCOVA) was run with garden-path comprehension accuracy as the dependent variable, RDI group as the independent variable and average, standardized WM score (average z-score for the three measures) and mean accuracy for positive lures in the 2-back task covariates. All outliers in the behavioral data (±2.5 SDs from the sample mean) were excluded from the analysis.

Results

Behavioral Data

Accuracy for garden-path sentences (M = 68.3%, SD = 14.4) was significantly lower than control sentences (M = 73.6%, SD = 11.4; $F(1,36) = 6.13$, $p < .05$, $\eta_p^2 = .15$).

Mean accuracy for the operation span, reading span, and symmetry span was 52.7 (SD = 15.9), 40.0 (SD = 15.9), and 21.1 (SD = 11.4), respectively. Analysis of the complex span tasks showed significant correlations between all three (all *r*s > .35, *p*s < .05, n = 35), as was expected (Kane, Conway, Miura & Colflesh, 2007). Mean accuracy on n+1 and n+2 lures in the 2-back task was 65.3% (SD = 27.6) and 75.9% (SD = 28.0), respectively. There were no significant correlations between accuracy in the lure conditions and complex span performance. This lack of correlation is consistent with previous findings (Kane, et al., 2007; Unsworth, Schrock, & Engle, 2004) and supports the finding that *n*-back and complex span tasks are indexing separate attentional mechanisms.

ERP Data

Garden-path sentences (compared to controls) showed a significant interaction of Type and AP over midline sites in the 500-700 and 700-900 ms time windows ($F(1,36) = 4.18$, $p < .05$, $\eta_p^2 = .10$ and $F(1,36) = 6.02$, $p < .05$, $\eta_p^2 = .14$, respectively) such that garden-paths elicited greater positivity than control sentences over posterior sites. Simple comparisons showed significant effects of type in posterior areas in both the 500-700 ($F(1,36) = 4.14$, $p < .05$, $\eta_p^2 = .10$) and 700-900 time windows ($F(1,36) = 4.93$, $p < .05$, $\eta_p^2 = .12$). There were no effects in the anterior sites.

RDI Analysis

Analysis of N400 and P600 effect magnitudes for garden path sentences showed a strong negative correlation ($r(32) = -.90$; $p < .001$; see Figure 1). The data suggest a continuum between strongly N400 and P600 dominant. Participants were divided into groups based on RDI values (negative values indicating N400 dominance and positive indicating P600 dominance). A total of 18 participants were N400 dominant and 19 were P600 dominant.

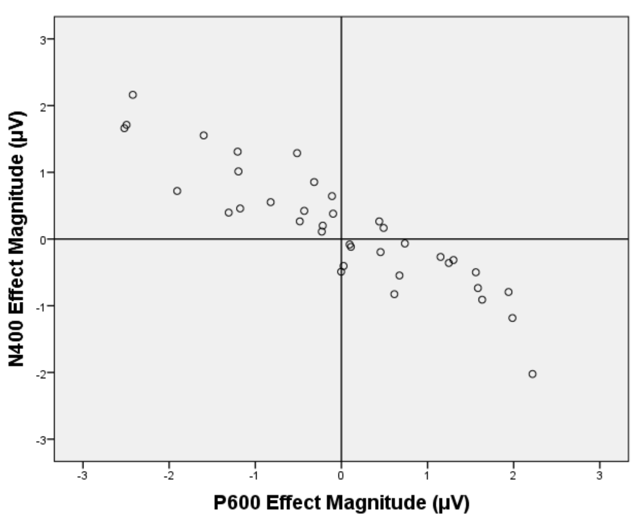

Figure 1. Scatter Plot of P600 and N400 effect magnitudes.

The effect of sentence type for each group was examined in three ERP data windows: 300-500, 500-700, and 700-900 ms. In N400 dominant participants there was a main effect of Type over dorsal areas in the 300-500 ($F(1,17) = 32.0$, $p < .001$, $\eta_p^2 = .65$), the 500-700 ($F(1,17) = 37.2$, $p < .001$, $\eta_p^2 = .69$), and the 700-900 ms windows ($F(1,17) = 27.4$, $p < .001$, $\eta_p^2 = .62$), such that garden-paths were more negative. There were also interactions of Type and AP in the 500-700 ($F(1,17) = 4.42$, $p = .05$, $\eta_p^2 = .21$) and the 700-900 ms windows ($F(1,17) = 5.87$, $p < .05$, $\eta_p^2 = .26$) with greater negativity at posterior sites.

For P600 dominant participants there was a main effect of Type over dorsal sites for all windows: 300-500 ($F(1,18) = 7.60$, $p < .05$, $\eta_p^2 = .30$), 500-700 ($F(1,18) = 19.2$, $p < .001$, $\eta_p^2 = .52$) and 700-900 ($F(1,18) = 21.2$, $p < .001$, $\eta_p^2 = .54$), such that garden-paths were more positive. Over midline sites there was a main effect of Type in the 500-700 ms time window ($F(1,18) = 4.42$, $p = .05$, $\eta_p^2 = .20$) and a significant interaction of Type and AP in the 700-900 ms window ($F(1,18) = 4.82$, $p < .05$, $\eta_p^2 = .21$), such that garden-paths elicited greater positivity over posterior sites.

A mixed factors ANOVA on the comprehension accuracy data with Type as a within participants factor and RDI group as a between participants factor showed a main effect of Type ($F(1,35) = 7.02$, $p < .05$, $\eta_p^2 = .17$) and an interaction of Type and group ($F(1,36) = 4.50$, $p < .05$, $\eta_p^2 = .11$) (see Figure 2). The interaction was driven by a simple effect of Type in the N400 dominant group only ($F(1,17) = 7.56$, $p < .05$, $\eta_p^2 = .31$).

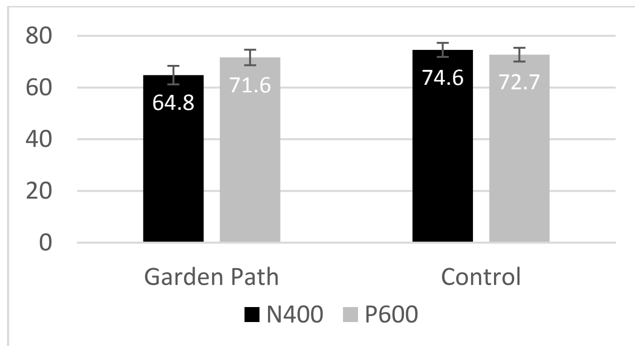

Figure 2. Mean comprehension accuracy for the two response dominance groups.

The fact that the P600 dominant group showed reduced garden-path effects in the accuracy data strongly suggests that the individual differences factor captured by the RDI is WM, or at least strongly related to WM, as individuals with greater WMC often show reduced garden-path effects in comprehension (Friederici et al., 1998). This question is addressed in the ANCOVA; the method was discussed prior in the data analysis section.

Prior to running the ANCOVA, it was necessary to determine that the covariates affected the dependent variable equally across the two groups. In the entire sample, there was a significant correlation between average complex span score and garden-path comprehension accuracy ($r(32) = .52$, $p < .01$), with each RDI group showing positive correlations (N400 dominant, $r(14) = .63$; P600 dominant, $r(16) = .48$) between the two variables. Using a Fisher transformation (Fisher, 1915), the difference between the group correlations was not significant ($z = .6$, $p > .50$). Similarly, mean positive 2-back lure accuracy correlated positively with garden-path comprehension accuracy ($r(33) = .38$, $p < .05$). Each group showed positive correlations (N400 dominant, $r(15) = .42$; P600 dominant, $r(16) = .35$), but the difference was not significant ($z = .2$, $p > .4$).

The ANCOVA showed a significant effect of response dominance on GP comprehension accuracy after controlling for RDI ($F(1,29) = 4.21$, $p < .05$, $\eta_p^2 = .13$) such that the P600 dominant group had greater accuracy. Complex span performance accounted for a significant amount of variance

($F(1,29) = 13.6$, $p < .05$, $\eta_p^2 = .32$), as did accuracy for positive 2-back lures ($F(1,29) = 5.28$, $p < .05$, $\eta_p^2 = .15$).

Discussion

The current study found evidence of distinct neural response profiles which were not apparent in the grand averaged data in neurologically normal, native English speakers during the processing of garden-path sentences and this individual differences measure predicted comprehension performance.

As in Tanner and Van Hell (2014), the N400 and P600 effect magnitudes fell into a continuous distribution and had a significant negative correlation ($r = -.90$). This correlation was greater than those found by Tanner and Van Hell for agreement and tense violations (*r*s = -.59 & -.60, respectively). An effect comparison using the Fisher transformation reveals that the correlation found in the current study was significantly greater than those in Tanner & Van Hell (2014) (all *p*s < .05). This discrepancy may suggest that response dominance is more pronounced or polarized in processing well-formed but challenging sentences (observed here) than that associated with syntactic violations (observed in Tanner & Van Hell).

The grand average data showed a standard P600 effect (increased positivity over posterior midline sites between 500 and 900 ms for the disambiguating word in garden-path sentences, in absence of early effects) which is consistent with the extant literature (Gouvea et al., 2010). When participants were divided into groups according to response dominance, N400 dominant individuals showed initial increased negativity followed by posterior negativity. P600 dominants showed a similar, but instead positive, pattern with increased positivity followed by later posterior positivities. In contrast to Tanner and Van Hell (2014), there were no clear topographical distinctions in the ERPs elicited by N400 and P600 individuals. The posterior effects in the N400 dominant group are in dorsal areas while in the P600 group they are in the midline. This does represent a topographical difference but, given that participants were sorted based on effect polarity over posterior dorsal and midline areas, it is not possible to make conclusions about qualitative differences in sentence type effects. The analysis of each group's ERPs confirms what was indicated in the correlational analysis: roughly half of the participants showed positivities while half showed negativities.

The ERPs do, however, shed light on the previously described instances of variability associated with garden-path resolution. Lack of P600 effects (Mecklinger, et al. 1995) and variable early effects (Friederici, et al., 2001; Gouvea, et al., 2010; Horberg, et al., 2013; Matzke, et al., 2002) could be the result of differing proportions of N400 and P600 dominant participants in the sample. In the current sample, 51% of participants were P600 dominant and 49% were N400 dominant and there were no early effects in the grand average analysis.

The key finding of the current study is the effect of response dominance on behavioral performance. Response dominance emerged as an effective predictor of comprehension accuracy such that P600 dominant participants had better comprehension accuracy for garden-path sentences. The results of the ANCOVA show that response dominance is not a proxy for WMC but rather a distinct individual difference measure. While WMC as assessed with both complex span and the *n*-back tasks, two tasks known to tap different aspects of attention control (Kane, et al., 2007; Unsworth et al., 2004), did contribute towards the successful recovery from temporary syntactic ability, this contribution was distinct from that of response dominance. This suggests that cognitive capacity alone does not limit the individual's ability to resolve garden-paths. Response dominance may, instead, indicate the engagement of specific parsing strategies.

Due to the ERP effect durations, the negativities in the N400 group are not canonical N400s, nor are the P600 group's positivities canonical P600s, thus interpreting them as such requires caution. That being said, the N400 is mostly associated with lexico-semantic processing and prediction (Kutas, Van Petten & Kluender 2006; Kutas & Federmeier, 2011) while the P600 is associated with combinatorial, syntactic processing (Gouvea, et al. 2010; O'Rourke & Van Petten, 2011). Applying Tanner and Van Hell's (2014) account for morphosyntactic violations to the current garden-path results, dominance of late posterior positivity may indicate an increased reliance on the combinatorial processing during parsing and, therefore, P600 dominant individuals would be in a better position to accurately reanalyze syntactic structure. In contrast, N400 dominant individuals may rely more on word-based predictions while parsing sentences, creating shallow syntactic representations. Failure to incorporate combinatorial syntactic information, thus, leads to a reduced ability to restructure the initial incorrect parse and increased likelihood of arriving at an incorrect interpretation.

The results of the current study extend the utility of Tanner and Van Hell (2014)'s RDI as an individual difference to a different syntactic context which elicits the P600 (i.e. garden-path sentences) showing that individuals in the sample exhibited distinct response profiles (either N400 or P600 dominant). Furthermore, response dominance during online sentence processing predicted offline comprehension and this relationship could not be reduced to individual differences in two aspects of WMC (storage/recall in the face of ongoing processing and discarding/suppression of active but no longer relevant information). While future research will reveal the neurocognitive underpinnings of response dominance, the findings of the current study establish this individual difference measure as a means of predicting behavior from neural activity.

References

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 Bull. Rev.* 12, 769-786.

Daneman, M., & Carpenter, P.A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, 19, 450–466.
Ferreira, F., Bailey, K. G., & Ferraro, V. (2002). Good-enough representations in language comprehension. *Current Directions in Psychological Science*, *11*(1), 11-15.
Fisher, R.A. (1915). "Frequency distribution of the values of the correlation coefficient in samples of an indefinitely large population". *Biometrika*, 10, 507–521
Friederici, A.D., Mecklinger, A., Spencer, K.M., Steinhauer, K. & Donchin, E. (2001). Syntactic parsing preferences and their on-line revisions: a spatio-temporal analysis of event-related brain potentials. *Cognitive Brain Research, 11,* 305-323.
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.
Frisch, S., Schlesewsky, M., Saddy, D. & Alpermann. (2002). The P600 as an indicator of syntactic ambiguity. *Cognition*, 85, B83-B92.
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.
Horberg, T., Koptjevskaja-Tamm, M. & Kallioinen, P. (2013). The neurophysiological correlate to grammatical function reanalysis in Swedish. *Language and Cognitive Processes. 28*, 388-416.
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.
Kaan, E., Swaab, T.Y.Y., 2003. Repair, revision, and complexity in syntactic analysis: an electrophysiological differentiation. *Journal Cognitive Neuroscience*, 15, 98–110.
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. J. Exp. Psychology Learning, Mem. Cogn. 33, 615-622.
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*, 621-647.
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.
Matzke, M., Mai, Heinke, Nager, W., Russeler, J. & Munte, T. (2002). The costs of freedom: and ERP study of non-canonical sentences. *Clinical Neuropsychology, 113*, 844-852.
Mecklinger, A., Schriefers, H., Steinhauer, K. & Friederici, A. (1995). Processing relative clauses varying on syntactic and semantic dimensions: An analysis with event-related potentials. *Memory & Cognition*, 23, 477-494.
O'Rourke, P. & Colflesh, G. J. H. (2014). Conflict resolution ability underpins the processing of temporary syntactic ambiguity: Evidence from ERPs. In P. Bello, M. Guarini, M. McShane, & B. Scassellati (Eds.), *Proceedings of the 36th Annual Conference of the Cognitive Science Society* (pp. 2056-2061). Austin, TX: Cognitive Science Society.
O'Rourke, P. L., & Van Petten, C. (2011). Morphological agreement at a distance: dissociation between early and late components of the event-related brain potential. *Brain research*, *1392*, 62-79.
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.
Slattery, T. J., Sturt, P., Christianson, K., Yoshida, M., & Ferreira, F. (2013). Lingering misinterpretations of garden path sentences arise from competing syntactic representations. *Journal of Memory and Language*, *69*(2), 104-120.
Tanner, D. & Van Hell, J.G. (2014). ERPs reveal individual differences in morphosyntactic processing. *Neuropsychologia, 56*, 289-301.
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.
Unsworth, N., Redick, T.S., Heitz, R.P., Broadway, J.M., & Engle, R.W. (2009). Complex working memory span tasks and higher-order cognition: A lantent-variable analysis of the relationship between processing and storage. *Memory*, *17*, 635-654.
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 52ndAvenue, College Park, MD 20742 (e-mail: porourke@casl.umd.edu). Special thanks to Mike Bunting and Valerie Karuzis.