

1
2
3 Running head: Brain-based Individual Differences
4
5
6

7 Brain-based Individual Differences in On-line L2 Grammatical Comprehension*
8
9

10 Darren Tanner
11 Department of Linguistics
12 University of Illinois at Urbana-Champaign
13
14

15 Kayo Inoue
16 Integrated Brain Imaging Center
17 Department of Radiology
18 School of Medicine
19 University of Washington
20
21

22 Lee Osterhout
23 Department of Psychology
24 University of Washington
25
26
27

28 *Acknowledgements

29 This research was supported by NIDCD grant R01DC01947 to Lee Osterhout and NSF BCS-
30 0951595 to Lee Osterhout and Darren Tanner. Darren Tanner also received support from the
31 William Orr Dingwall Neurolinguistics Dissertation Fellowship and NSF OISE-0968369. We
32 would like to thank Julia Herschensohn, Judith McLaughlin, Janet Nicol, Janet Van Hell and
33 Eleonora Rossi for thoughtful comments and discussion. Our thanks also go to the participants,
34 as well as Geoff Valentine, Kristie Fisher, Shota Moma, and Elliot Collins for their help in
35 acquiring the data. We would also like to thank two anonymous reviewers for helpful comments
36 on a previous version of this manuscript. Any remaining errors are our own.
37
38
39

40 Address for correspondence:

41 Darren Tanner
42 707 S Matthews St
43 4080 Foreign Languages Building
44 MC 168
45 Urbana, IL 61801 USA
46 Email: dstanner@gmail.com
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Abstract

Using event-related potentials, we investigated the impact of a range of individual difference measures related to L2 learning on proficient L1 Spanish – L2 English bilinguals' brain responses during L2 morphosyntactic processing. Although grand-mean ERP analyses revealed a biphasic N400-P600 response to English subject-verb agreement violations, subsequent analyses showed that participants' brain responses varied along a continuum between N400 and P600 dominance. To investigate this pattern, we introduce two novel ERP measures which independently quantify relative brain response type and overall magnitude. Multivariate analyses revealed that larger overall brain responses were associated with higher L2 proficiency, while relative brain response type (N400 or P600) was predicted by a coalition of variables, most notably learners' motivation and age of arrival in an L2 environment. Our findings show that aspects of a learner's background can differentially impact a learner's overall sensitivity to L2 morphosyntax and qualitative use of linguistic cues during processing.

Keywords: ERP, N400, P600, individual differences, morphosyntax, second language acquisition

Brain-based Individual Differences in On-line L2 Grammatical Comprehension

In an interconnected global society, knowledge of a second language (L2) is an increasingly indispensable skill. Growing globalization in trade, education, and politics requires large numbers of individuals with strong L2 skills, and increasing international immigration is creating a large population of individuals who find themselves needing to master a nonnative language rapidly. However, there is a great deal of variability in the rate, style, trajectory, and ultimate success of L2 learning in adulthood. Understanding what individual-level factors are associated with variability in L2 learning and comprehension is a fundamental question both for cognitive scientists interested in language learning and plasticity in general, as well as for applied researchers interested in identifying cognitive skills or strategies underlying successful learning, identifying gifted language learners for selection in language training programs, or more generally improving learner outcomes. Indeed, some research has sought to characterize predictive cognitive factors or learner strategies associated with rapid and successful learning (e.g., Carroll, 1962; Dörnyei, 2005; Naiman, Fröhlich, Stern, & Todesco, 1996; Skehan, 1989). Other studies focusing on experiential factors associated with individuals' long-term learning outcomes have shown that greater success is associated with early immersion, higher motivation to learn, and more frequent L2 use in daily life (e.g., Birdsong & Molis, 2001; Dörnyei, 2005; Flege, Yeni-Komshian, & Liu, 1999; Gardner, Tremblay, & Masgoret, 1997; Johnson & Newport, 1989).

These studies have largely relied on offline measures such as learners' performance on large test batteries, or subjective measures such as teachers' ratings of learner development over time. However, classroom evaluations and pen-and-paper tests represent a limited way of

1
2
3 understanding L2 learning. Proficiency testing provides a continuous, but ultimately
4
5 unidimensional outcome measure of language knowledge. Moreover, the results of these off-line
6
7 measures may not generalize to learners' real-time language use in context. On the other hand,
8
9 studying event-related brain potentials (ERPs) allows researchers to investigate the neural
10
11 mechanisms of real-time language comprehension with millisecond-level temporal resolution.
12
13 ERPs reflect individuals' brainwave activity that is time- and phase-locked to the presentation of
14
15 stimuli, such as words in sentences. Additionally, ERPs are multidimensional, as they provide
16
17 both quantitative (e.g., effect onset timing and amplitude) and qualitative (e.g., positive or
18
19 negative effect polarity and scalp distribution) information about processing mechanisms
20
21 underlying language comprehension (see e.g., Handy, 2005; Luck, 2005, for thorough
22
23 introductions to ERPs). For example, ERP studies of native language (L1) sentence processing
24
25 have consistently shown a neurocognitive dissociation between the processing of lexico-semantic
26
27 and morphosyntactic anomalies. Lexico-semantic manipulations typically trigger an enhanced
28
29 negativity, peaking around 400ms poststimulus (the N400 effect: Bentin, 1987; Kutas & Hillyard,
30
31 1980, 1984; Osterhout & Nicol, 1999; though see e.g., Kim & Osterhout, 2005; Kuperberg,
32
33 Kreher, Tatiana Sitnikova, Caplan, & Holcomb, 2007; van de Meerendonk, Kolk, Vissers, &
34
35 Chwilla, 2010), whereas a range of syntactic manipulations, including violations of
36
37 morphosyntactic rules, typically elicit a large positive-going wave with a maximum around
38
39 600ms poststimulus (the P600 effect: Friederici, Hahne, & Mecklinger, 1996; Kaan & Swaab,
40
41 2003; Osterhout & Holcomb, 1992; Osterhout & Mobley, 1995).¹
42
43
44
45
46
47
48
49

50
51 Based on these L1 findings, N400 and P600 effects can be seen as indices of two
52
53 independent, but highly interactive 'streams' of processing which are differentially sensitive to
54
55 linguistic cues. Recent conceptions of the N400/P600 dichotomy have posited that dominance of
56
57
58
59
60

1
2
3 one response or the other reflects a competition between a lexically- or memory-based heuristic
4
5 processing stream on the one hand, and a combinatorial, algorithmic stream on the other (Kim &
6
7 Osterhout, 2005; Kuperberg, 2007; Osterhout, Kim, & Kuperberg, 2012; see e.g., van de
8
9 Meerendonk et al., 2010; Van Herten, Chwilla, & Kolk, 2006, for a related view). In the context
10
11 of sentence processing, increased N400 amplitudes are seen when a sentential or discourse
12
13 context makes a given lexical item more difficult to access from long-term semantic memory or
14
15 integrate into the still-being-constructed semantic or discourse representation (Van Berkum,
16
17 Hagoort, & Brown, 1999; Hagoort, Hald, Bastiaansen, & Petersson, 2004; Kuperberg, 2007;
18
19 Kutas & Federmeier, 2000; Lau, Phillips, & Poeppel, 2008). P600 effects, however, are elicited
20
21 by engagement of combinatorial processes, which frequently rely on linguistic constraints. These
22
23 constraints include morphosyntactic rules or predictions (Hagoort & Brown, 2000; Hahne &
24
25 Friederici, 1999; Osterhout & Mobley, 1995; Osterhout & Nicol, 1999) and verb-argument
26
27 combinatorial constraints, such as animacy restrictions (Kuperberg, Caplan, Tatiana Sitnikova,
28
29 Eddy, & Holcomb, 2006; Kuperberg et al., 2007; Paczynski & Kuperberg, 2011). Importantly,
30
31 this response dichotomy can be used to test whether encountering a particular type of anomaly
32
33 preferentially engages memory-based lexical or algorithmic combinatorial processing
34
35 mechanisms.
36
37
38
39
40
41
42

43
44 Given their unique sensitivity to different levels of processing, ERPs can also be useful in
45
46 characterizing the cognitive mechanisms underlying L2 comprehension. Over the last decade
47
48 there has been an enormous growth of interest in the neurocognitive substrates of L2 learning
49
50 and processing (see Van Hell & Tokowicz, 2010; McLaughlin et al., 2010; Osterhout,
51
52 McLaughlin, Pitkänen, Frenck-Mestre, & Molinaro, 2006; Steinhauer, White, & Drury, 2009 for
53
54 reviews of L2 ERP research), and much of this research has focused on identifying whether the
55
56
57
58
59
60

1
2
3 neural mechanisms underlying L2 comprehension are fundamentally different from or similar to
4 those observed in L1 populations. As much of the work on L1 processing has assumed that
5 monolinguals always show a P600 effect to syntactic violations (though see below for important
6 caveats to this generalization), one of the driving questions in L2 research on morphosyntactic
7 processing is therefore whether learners can similarly show P600 effects to L2 syntactic
8 violations, and if so, at what point in acquisition they may show them. Although some studies
9 have shown that L2 syntactic violations elicit N400 effects in learners in the earliest stages of
10 acquisition (McLaughlin et al., 2010; Osterhout et al., 2006; Tanner, McLaughlin, Herschensohn,
11 & Osterhout, 2013), P600 effects have been observed in relatively low proficiency L2 learners
12 processing violations of syntactic rules common to the L1 and L2 (e.g., McLaughlin et al., 2010;
13 Rossi, Gugler, Friederici, & Hahne, 2006; Tokowicz & MacWhinney, 2005), as well as high
14 proficiency learners processing novel L2 features (e.g., Frenck-Mestre, Foucart, Carrasco, &
15 Herschensohn, 2009; Gillon Dowens, Guo, Guo, Barber, & Carreiras, 2011; Gillon Dowens,
16 Vergara, Barber, & Carreiras, 2010; Morgan-Short, Sanz, Steinhauer, & Ullman, 2010).
17
18 Nonetheless, there are some exceptions to this generalization, where P600 effects have failed to
19 be found even in relatively proficient learner groups, usually when L2 features are not realized in
20 the L1 or have different morphological instantiations (e.g., Chen, Shu, Liu, Zhao, & Li, 2007;
21 Ojima, Nakata, & Kakigi, 2005; Sabourin & Haverkort, 2003; Sabourin & Stowe, 2008).
22
23

24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46 ERPs' differential sensitivity to levels of linguistic processing makes them an ideal
47 candidate for studying how individual difference measures map on to variation in individuals'
48 ERP responses. However, very little research has taken this approach in either L1 or L2
49 processing. For example, standard approaches to the quantification of ERPs treat inter-subject
50 variability as a source of noise in statistical analyses (e.g., in the error term in ANOVA statistics),
51
52
53
54
55
56
57
58
59
60

1
2
3 and most published ERP waveforms represent the central tendency after averaging the raw
4
5 electroencephalogram across both trials and subjects. These waveforms therefore may not
6
7 accurately depict any individual's brain response to a stimulus on a particular trial. Additionally,
8
9 ERPs' multidimensional nature may make quantifying the relevant dimension of variation
10
11 difficult, as individuals' effects may differ in amplitude, polarity, timing, or all three. A limited
12
13 number of studies have investigated the impact of individual differences using grouped designs,
14
15 where groups are determined by splits on some relevant background measure (e.g., working
16
17 memory span, comprehension performance, L2 proficiency, or age of arrival: King & Kutas,
18
19 1995; Ojima et al., 2005; Rossi et al., 2006; Vos, Gunter, Kolk, & Mulder, 2001; Weber-Fox &
20
21 Neville, 1996; Weckerly & Kutas, 1999). However, this approach provides little information as
22
23 to the nature of the relationship between the background and outcome measures (e.g., ERP
24
25 amplitude or polarity), which may be linear and graded (see Van Hell & Tanner, 2012, for
26
27 related discussion). On the other hand, multivariate correlation- and regression-based statistics
28
29 have the power to more accurately capture the continuous nature of individual variation, but have
30
31 only recently been used to study individual differences in ERPs (e.g., Bond, Gabriele, Fiorentino,
32
33 & Alemán Bañón, 2011; Moreno & Kutas, 2005; Newman, Tremblay, Nichols, Neville, &
34
35 Ullman, 2012; Ojima, Matsuba-Kurita, Nakamura, Hoshino, & Hagiwara, 2011; Pakulak &
36
37 Neville, 2010; Tanner et al., 2013).

38
39 Our goal here was to investigate the impact of a range of individual difference measures
40
41 related to L2 learning on ERP responses to morphosyntactic anomalies in late, but highly
42
43 proficient bilinguals, specifically using a multivariate approach. For example, what factors
44
45 predict relative reliance on memory-based processes versus combinatorial analyses, and what
46
47 predicts the magnitude of the relevant effects? One particular individual difference variable of
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 interest in recent language processing studies has been proficiency (e.g., Hopp, 2006, 2010;
4 Jackson & Van Hell, 2011; Newman et al., 2012; Ojima et al., 2005; Pakulak & Neville, 2010;
5
6 Rossi et al., 2006; Van Hell & Tanner, 2012). In a large-scale review paper, Steinhauer, White,
7
8 and Drury (2009) present a proficiency-based neurocognitive model of L2 development. They
9
10 propose that low-proficiency learners may show N400 effects to grammatical violations (in line
11
12 with longitudinal and cross-sectional studies of early-stage L2 learners: McLaughlin et al., 2010;
13
14 Osterhout et al., 2006; Tanner et al., 2013), but that given high enough proficiency, learners will
15
16 show large P600 or biphasic LAN-P600 responses, as are assumed to be consistently elicited in
17
18 native speakers (see also Ullman, 2001, 2005, for a similar proposal). Based on this model, one
19
20 prediction is that marked inter-subject variability in the type or size of brain response may be
21
22 present in early-stage learners (e.g., Tanner et al., 2013), but this variability should decrease at
23
24 high proficiency as learners' brain responses approximate native-like targets. That is, variability
25
26 in highly proficient bilinguals' brain response type or P600 magnitude should be trivial, and to
27
28 the extent it exists, should be related to L2 proficiency level.
29
30
31
32
33
34
35

36
37 However, some recent findings from the L1 processing literature suggest that variability
38
39 exists among even proficient language users, including monolinguals processing their native
40
41 language – a group which has traditionally been assumed to be relatively homogenous in the
42
43 ERP literature. In some cases this variability was related to participants' L1 proficiency,
44
45 supporting Steinhauer et al.'s claims. For example, Newman and colleagues (Newman et al.,
46
47 2012) showed that both L1 and L2 users' N400 magnitudes to semantic anomalies correlated
48
49 with proficiency measures. Pakulak and Neville (2010) showed that the laterality of an early
50
51 negativity and the magnitude of the P600 effect in response to simple English phrase structure
52
53 violations varied continuously with respect to monolingual English speakers' L1 proficiency.
54
55
56
57
58
59
60

1
2
3 However, proficiency is not always implicated in individuals' processing profiles. Osterhout
4 (1997) showed that, among highly literate university students, certain types of syntactic
5 anomalies elicited P600s in some individuals, but N400s in others. Oines, Miyake and Kim
6 (2012) showed that, after controlling for individuals' language experience, vocabulary, and
7 spatial working memory, verbal working memory measures reliably predicted whether
8 implausible verb-argument relations elicited N400 or P600 effects: those with larger spans
9 showed relatively larger P600s, while those with smaller spans showed relatively larger N400s
10 (see also Nakano, Saron, & Swaab, 2010, for similar findings). Overall, this suggests that there
11 are robust individual differences in the size and type of ERP responses among monolinguals
12 processing their L1, such that we might expect to see variability in brain response type and size
13 even among very proficient L2 speakers. It remains open to what extent proficiency is implicated
14 in these differences, however.

15
16 In addition to proficiency, a number of other individual difference measures have been
17 associated with L2 learning and processing profiles. For example, Weber-Fox & Neville (1996,
18 1999) suggest that age of arrival in an L2 environment may be a crucial determinant in whether
19 syntactic anomalies elicit P600 effects (see Birdsong & Molis, 2001; Johnson & Newport, 1989
20 for age effects on L2 behavior). Behavioral research has additionally implicated learner
21 motivation, amount of L2 exposure (e.g., length of residence in an L2 environment), and
22 frequency of L2 use in daily in determining learner profiles (e.g., Dörnyei & Skehan, 2003;
23 Dörnyei, 2005; Flege et al., 1999; Gardner, Tremblay, & Masgoret, 1997; Skehan, 1989), though
24 no ERP research has directly investigated potential impacts of these factors in determining how
25 information is processed on-line.

26
27 To investigate these questions, we recorded ERPs while late, but highly proficient L1
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 Spanish – L2 English bilinguals processed English sentences involving subject-verb agreement, a
4 morphosyntactic rule shared between Spanish and English. Agreement is well-studied using
5 ERPs in L1 processing (see Molinaro, Barber, & Carreiras, 2011, for a recent review), as well as
6 in novice and intermediate L2 learners of typologically similar languages (e.g., Frenck-Mestre,
7 Osterhout, McLaughlin, & Foucart, 2008; McLaughlin et al., 2010; Tanner et al., 2013) and
8 proficient L2 learners of typologically distinct languages (Chen et al., 2007; Ojima et al., 2005).
9
10 Based on previous research investigating the processing of features shared between the L1 and
11 L2 (e.g., Osterhout et al., 2006; Tokowicz & MacWhinney, 2005) we predicted that agreement
12 violations would elicit large P600 effects in our highly experienced learners. However, results
13 showed a great deal of variability in both the type and magnitude of learners' brain responses.
14
15 Using multiple regression, we investigated this variability by assessing the impact of several
16 factors which are either known from behavioral studies to impact learning outcomes or which
17 have been shown to impact language processing: age of arrival, length of residence, frequency of
18 L2 use, language proficiency, and learner motivation. In doing this, we introduce two new
19 outcome measures for ERP waveforms which independently quantify individuals' relative brain
20 response type (N400 or P600) and overall ERP response magnitude.
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40

41 **Method**

42 *Participants*

43
44 Our participants included 24 native Spanish speakers who had acquired English as an L2.
45
46 Data from four participants were excluded from final analysis due to excessive eye movement or
47 other artifact in the raw EEG. Thus, data from 20 participants (7 male) are reported here. All
48 participants were strongly right-handed as assessed by an abridged version of the Edinburgh
49 Handedness Inventory, and had normal or corrected-to-normal vision. In order to ensure that
50
51
52
53
54
55
56
57
58
59
60

1
2
3 participants had acquired the L2 after childhood, but had sufficient exposure to achieve high
4
5 proficiency, participants were screened such that they had not been exposed to English in the
6
7 home, had first moved to an English-speaking country at age 15 or later, and had lived immersed
8
9 in an English-speaking environment for a minimum of five years. Participants completed a
10
11 language background questionnaire, which included self-reports on age of initial exposure to
12
13 English (AoE), age of arrival in an English-speaking environment (AoA), and total length of
14
15 residence in an English-speaking environment (LoR), as well as proficiency self-ratings for their
16
17 L1 Spanish and L2 English on a Likert-scale between 1 (no proficiency) and 7 (perfect
18
19 proficiency). Other questions asked participants about their frequency of use of English in
20
21 various contexts in daily life, an overall estimate of English use between 1 (never use English)
22
23 and 7 (always use English), as well as their motivation to speak English like a native speaker
24
25 between 1 (not important to sound like a native speaker at all) and 7 (extremely important to
26
27 sound like a native speaker). Participants also completed a pen-and-paper proficiency test
28
29 consisting of 50 questions selected from the Michigan Examination for the Certificate of
30
31 Proficiency in English (ECPE). Participants' responses are reported in Table 1. Additionally,
32
33 eight of the participants reported no significant competency in languages other than Spanish and
34
35 English. The remaining participants reported varying non-native competence in other European
36
37 languages, including French, Italian, German, Portuguese, and Catalan. Participants provided
38
39 informed consent and received a small amount of cash for taking part.
40
41
42
43
44
45
46
47

48 <Insert Table 1 about here>
49

50 ***Materials***

51
52
53 Stimuli were sentences that were either grammatically correct or contained a violation of
54
55 subject-verb agreement. All sentences contained a singular subject noun, followed by a
56
57
58
59
60

1
2
3 prepositional phrase modifier containing another singular noun, followed by a verb that either
4
5 agreed or disagreed with the subject noun in number (*is/are, was/were, or has/have*), followed by
6
7 a short predicate (e.g., *The winner of the big trophy has/*have proud parents*). Two hundred
8
9 forty sentence frames were constructed with eight versions of each sentence. Two versions
10
11 corresponded to the grammatical-ungrammatical sentence pairs reported here; the other six
12
13 versions investigated agreement in more complex syntactic configurations (see Tanner, 2011;
14
15 Tanner, Nicol, Herschensohn, & Osterhout, 2012). Sentences were distributed across eight
16
17 experimental lists in a Latin square design, such that each listed contained only one version of
18
19 each sentence frame. Each list thus contained 30 grammatical and 30 ungrammatical sentences in
20
21 the experiment reported here. Experimental sentences were pseudo-randomized among other
22
23 sentences belonging to other experiments not reported here. Half of the filler sentences were
24
25 anomalous, either with an anomaly of verb tense (e.g., **The man will cooked the food in the*
26
27 *refrigerator*) or lexical semantics (e.g., **Jane will bake a book in her spare time*). Each list
28
29 contained 540 sentences, half of which contained either a syntactic or semantic anomaly.
30
31
32
33
34
35

36 ***Procedure***

37
38 Participants took part in three sessions. The first session involved completion of the
39
40 background questionnaires and proficiency test; ERPs were recorded during the two subsequent
41
42 sessions. Because of the large number of experimental sentences, participants saw half of the
43
44 sentences at each session; experimental sentences were balanced across sessions, such that each
45
46 participant saw an equal number of sentences from each condition at both session. ERPs were
47
48 first averaged within sessions. No between-session differences were found, so ERPs were
49
50 subsequently averaged across sessions. All reported results reflect these cross-session averages.
51
52
53
54
55 During ERP recording participants were seated in a comfortable recliner in front of a CRT
56
57
58
59
60

1
2
3 monitor. Participants were instructed to relax and minimize movements and eye blinks while
4
5 reading and to read each sentence as normally as possible. Each trial consisted of the following
6
7 events: each sentence was preceded by a blank screen for 1000ms, followed by a fixation cross,
8
9 followed by a stimulus sentence, presented one word at a time. The fixation cross and each word
10
11 appeared on the screen for 400ms followed by a 200ms ISI. Sentence-ending words appeared
12
13 with a full stop. This screen was followed by a “yes/no” prompt asking for a sentence
14
15 acceptability judgment. Participants were instructed to respond “yes” to sentences that were
16
17 well-formed and semantically coherent and “no” to sentences that were ungrammatical or
18
19 semantically incoherent. Participants were randomly assigned to use either their left or right hand
20
21 for the “yes” response.
22
23
24
25
26

27 *Data Acquisition and Analysis*

28
29 Continuous EEG was recorded from 19 tin electrodes attached to an elastic cap (Electro-
30
31 cap International) in accordance with the 10-20 system (Jasper, 1958). Eye movements and
32
33 blinks were monitored by two electrodes, one placed beneath the left eye and one placed to the
34
35 right of the right eye. Electrodes were referenced to an electrode placed over the left mastoid.
36
37 EEG was also recorded from an electrode placed on the right mastoid to determine if there were
38
39 experimental effects detectable on the mastoids. No such effects were found. EEG signals were
40
41 amplified with a bandpass of 0.01-100 Hz (3db cutoff) by an SAI bioamplifier system. ERP
42
43 waveforms were filtered off-line below 15 Hz. Impedances at scalp and mastoid electrodes were
44
45 held below 5 k Ω and below 15 k Ω at eye electrodes.
46
47
48
49

50
51 Continuous analog-to-digital conversion of the EEG and stimulus trigger codes was
52
53 performed at a sampling frequency of 250 Hz. ERPs, time-locked to the onset of the critical verb
54
55 (underlined in the examples above), were averaged off-line for each participant at each electrode
56
57
58
59
60

1
2
3 site in each condition. Trials characterized by eye blinks, excessive muscle artifact, or amplifier
4 blocking were not included in the averages; 5.8% and 5.5% of grammatical and ungrammatical
5 trials were rejected, respectively. ERPs were quantified as mean amplitude within a given time
6 window. All artifact-free trials were included in the ERP analyses. Because of a small amount of
7 noise in the prestimulus interval, a 50ms prestimulus to 50ms poststimulus baseline was used to
8 mitigate early differences in the waveforms. In accordance with previous literature and visual
9 inspection of the grand mean waveforms, the time windows 400-500ms² and 500-1000ms were
10 chosen, as they correspond roughly to the N400 and P600 effects, respectively. For grand mean
11 analyses, ANOVAs were calculated within each time window with grammaticality (grammatical,
12 ungrammatical) as a within-subjects factor. Data from midline (Fz, Cz, Pz), medial-lateral (right
13 hemisphere: Fp2, F4, C4, P4, O2; left hemisphere: Fp1, F3, C3, P3, O1), and lateral-lateral (right
14 hemisphere: F8, T8, P8; left hemisphere: F7, T7, P7) electrode sites were treated separately in
15 order to identify topographic and hemispheric differences. ANOVAs on midline electrodes
16 included electrode as an additional within-subjects factor (3 levels), ANOVAs on medial-lateral
17 electrodes included hemisphere (2 levels) and electrode pair (5 levels) as additional within-
18 subjects factors, and ANOVAs over lateral-lateral electrodes included hemisphere (2 levels) and
19 electrode pair (3 levels) as additional within-subjects factors. The Greenhouse-Geisser correction
20 for inhomogeneity of variance was applied to all repeated measures on ERP data with greater
21 than one degree of freedom in the numerator. In such cases, the corrected *p*-value is reported.
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48 Regression models will be described with the results.
49

50 Results

51 *Grand Mean Results*

52
53
54
55 Results from the end-of-sentence judgment task showed that participants were highly
56
57
58
59
60

1
2
3 accurate in judging the acceptability of the sentences (grammatical mean proportion correct = .90,
4
5 SE = .02; ungrammatical mean proportion correct = .88, SE = .04). Grand mean ERP results for
6
7 grammatical and ungrammatical verbs are depicted in Figure 1. Visual inspection of the
8
9 waveforms showed that, relative to grammatical verbs, ungrammatical verbs elicited a small
10
11 biphasic waveform characterized by a centrally-distributed negativity between approximately
12
13 400-500ms poststimulus (N400), followed by a broadly-distributed positivity (P600). Statistical
14
15 analysis confirmed these observations. In the 400-500ms window there was an effect of
16
17 grammaticality that was strongest over central electrodes [grammaticality x electrode interaction:
18
19 midline, $F(2, 38) = 4.967, p < .03$; medial-lateral, $F(4, 76) = 4.721, p < .02$], and which showed a
20
21 stronger left-hemisphere distribution over lateral-lateral sites [grammaticality x hemisphere
22
23 interaction, $F(1, 19) = 9.190, p < .01$]. Between 500-1000ms the positivity reached significance
24
25 over a broad portion of the scalp [main effect of grammaticality: midline, $F(1, 19) = 8.719, p$
26
27 $< .01$; medial-lateral: $F(1, 19) = 10.292, p < .01$; lateral-lateral: $F(1, 19) = 7.519, p < .02$]. Over
28
29 lateral-lateral sites the positivity showed a stronger posterior [grammaticality x electrode
30
31 interaction: $F(2, 38) = 4.299, p < .03$] and right-hemisphere distribution [grammaticality x
32
33 hemisphere interaction: $F(1, 19) = 10.776, p < .01$].

34
35
36
37
38
39
40
41 <Insert Figure 1 about here>

42 43 ***Individual Differences Analyses***

44
45
46 Despite the significant biphasic N400-P600 results in the omnibus ANOVA, inspection
47
48 of individuals' ERP waveforms showed that some individuals showed primarily either an N400
49
50 or P600 to ungrammatical verbs, but not both, while others showed a biphasic response.
51
52 Following Inoue & Osterhout (2013) and Tanner et al. (2013), we regressed individuals' N400
53
54 effect magnitude onto their P600 effect magnitude. To compute these measures, we calculated
55
56
57
58
59
60

1
2
3 participants' mean amplitudes in a centro-parietal region of interest (electrodes C3, Cz, C4, P3,
4
5 Pz, P4), where N400 and P600 effects are typically largest. N400 effect magnitude was
6
7 calculated as mean activity in the grammatical minus ungrammatical condition between 400-
8
9 500ms; P600 effect magnitude was calculated as mean activity in the ungrammatical minus
10
11 grammatical conditions between 500-1000ms. Results showed the two effects to be highly
12
13 negatively correlated, $r = -.737$, $p = .0002$, Figure 2. Brain responses showed a continuous
14
15 distribution between negativity-dominance, biphasic, and positivity-dominance across
16
17 individuals; as one effect increased, the other decreased to a similar degree. To illustrate these
18
19 differences, we averaged ERPs separately for those who showed a negativity-dominance (e.g.,
20
21 those above/to the left of the dashed line in Figure 2) and those who showed a positivity
22
23 dominance (e.g., those below/to the right of the dashed line in Figure 2). Figure 3 shows mean
24
25 waveforms for these two groups. Those in the negativity-dominant group showed a large,
26
27 centrally-distributed negativity to ungrammatical verbs between approximately 400-600ms (an
28
29 N400), but no later positivity; those in the positivity-dominant group showed no earlier
30
31 negativity, but instead a large, posteriorly distributed positivity between 500-1000ms (a P600).
32
33 However, as can be seen in Figure 2, the true distribution of brain response dominance across
34
35 individuals is continuous. Thus, grand mean waveforms depicted an effect that was not
36
37 representative of most individuals' ERP profiles.
38
39
40
41
42
43
44

45
46 <Insert Figure 2 about here>
47

48
49 <Insert Figure 3 about here>
50

51 To investigate what factors predict the type and magnitude of ERP response to agreement
52
53 violations in this group of learners, we computed two new measures to serve as the dependent
54
55 variables (DVs) in multiple regression models. For the first measure, the Response Dominance
56
57
58
59
60

1
2
3 Index (RDI), we measured each individual's relative response dominance (N400 or P600) by
4 fitting the individual's least squares distance from the equal effect sizes line (the dashed line in
5 Figure 2) using perpendicular offsets. RDI values near zero reflect relatively equal-sized N400
6 and P600 effects, while more negative or positive values reflect relatively larger negativies or
7 positivities across both time windows, respectively. While the RDI gives an indication about
8 response dominance, it does not provide a strong measure of overall response size. For example,
9 it would not differentiate between two individuals who each showed equal-sized N400 and P600
10 effects, but where one individual showed relatively small effects and the other showed a large
11 biphasic response. Both individuals would have an RDI value near zero, though the second
12 individual is clearly more sensitive to the agreement violations than the first. We therefore
13 computed a second measure, the Response Magnitude Index (RMI), by calculating each
14 individual's Euclidian distance from zero in Figure 2. The RMI gives a measure of the overall
15 level of sensitivity an individual shows to the agreement violations within the N400 and P600
16 time windows. Larger values of the RMI indicate relatively greater neural responses to the
17 agreement violations across both time windows, regardless of the type of the response. Equations
18 (1) and (2) show how the RMI and RDI were computed, where N400 and P600 refer to mean
19 amplitude between 400-500ms and 500-1000ms, respectively, averaged within the centro-
20 parietal ROI (C3, Cz, C4, P3, Pz, P4).
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44

$$RMI = \sqrt{(N400_{\text{gram}} - N400_{\text{ungram}})^2 + (P600_{\text{ungram}} - P600_{\text{gram}})^2} \quad (1)$$

$$RDI = \frac{(P600_{\text{ungram}} - P600_{\text{gram}}) - (N400_{\text{gram}} - N400_{\text{ungram}})}{\sqrt{2}} \quad (2)$$

45
46
47
48
49
50
51
52
53
54
55
56 We fit two multiple regression models with the RMI and RDI serving as the respective
57
58
59
60

1
2
3 DVs. Five background measures known to affect learning outcomes and processing profiles were
4 selected as independent variables for the models: AoA, LoR, frequency of L2 use, L2
5 proficiency (ECPE scores), and motivation to speak like a native speaker (see Table 1). Because
6 of right-skewed distributions, AoA and LoR were log-transformed prior to entry into the
7 regression models. Additionally, since all participants reported having relatively high motivation
8 to speak like a native English speaker, this measure was dichotomized into those reporting very
9 high motivation (i.e., those reporting 7 on the Likert-scale, $n=11$) and those reporting high
10 motivation (i.e., those reporting 5 or 6 on the Likert-scale, $n=9$). ‘High’ motivation to speak like
11 a native speaker set as the reference level (0) and ‘very high’ motivation set to 1 in the models.
12 We also fit regression models with the motivation variable undichotomized. Results were
13 qualitatively similar. However, we chose to include the dichotomized variable in the final model
14 because the distribution of the variable was highly skewed and there were a large number of
15 observations at one end of the distribution. Dichotomization is an appropriate transformation in
16 such circumstances (MacCallum, Zhang, Preacher, & Rucker, 2002). Distributional statistics for
17 the DVs and background measures entered into the models are reported in Table 2; the
18 correlation matrix for the measures is reported in Table 3. Collinearity diagnostics showed that
19 multicollinearity among predictor variables was not a problem (tolerances $> .7$, VIFs < 1.5).
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Residuals of both models were approximately normally distributed.

<Insert Table 2 about here>

<Insert Table 3 about here>

The total model predicting the RMI was not significant ($R^2 = .371$, Adjusted $R^2 = .146$,
 $F(5, 14) = 1.650$, $p = .211$), though one individual predictor reached significance. After
controlling for the effects of AoA, LoR, frequency of use, and motivation to speak like a native

1
2
3 speaker, English proficiency showed a positive association with the RMI ($B = .209$, $\beta = .509$,
4
5 partial- $r = .517$, $p = .04$). This shows that those who were more proficient in an offline pen-and-
6
7 paper test also showed larger overall neural sensitivity to agreement violations. In order to
8
9 specifically test the prediction made by Steinhauer et al. (2009; see also Rossi et al., 2006) that
10
11 N400 magnitude would decrease and P600 magnitude would increase with increasing
12
13 proficiency, we correlated individuals' respective effect magnitudes with ECPE proficiency
14
15 measures. Neither N400 magnitudes ($r = .070$) nor P600 magnitudes ($r = .084$) correlated with
16
17 individuals' L2 proficiency, suggesting that proficiency is most associated with overall response
18
19 magnitude, as measured by the RMI, rather than the magnitude of specifically the N400 or P600
20
21 response in particular. The model predicting brain response dominance (the RDI) was highly
22
23 significant (Table 4), such that the linear combination of AoA, LoR, frequency of use, L2
24
25 proficiency and motivation to speak like a native speaker accounted for approximately 54% of
26
27 the variance in response dominance. Two individual predictors also reached significance. AoA
28
29 and motivation to speak like a native speaker each uniquely predicted response dominance: those
30
31 who experienced earlier immersion in an English-speaking environment and who reported higher
32
33 motivation to speak in a native-like way showed stronger P600-dominant brain responses, while
34
35 those who arrived later and reported lower motivation to speak like a native speaker were more
36
37 likely to show N400-dominant responses. To test whether the linear combination of the non-
38
39 significant predictor variables (LoR, proficiency, and frequency of use) together provided
40
41 predictive power over and above motivation and AoA alone, we subdivided the RDI model into
42
43 two hierarchically related models: the first (nested) model included only the two significant
44
45 predictors (AoA and motivation), while the second (full) model added the remaining three
46
47 predictors. In the first model, AoA and motivation to speak like a native speaker were
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 individually significant (partial-rs = -.603 and .703, respectively, $ps < .007$), and together
4
5 accounted for a significant amount of variance in the RDI, $R^2 = .533$, Adjusted $R^2 = .479$, $F(2,$
6
7 $17) = 9.718$, $p = .002$. Adding LoR, proficiency, and frequency of use did not add a significant
8
9 amount of predictive power to the model, however, R^2 -change = .127, $F(3, 14) = 1.737$, $p = .205$.
10
11 <Insert Table 4 about here>
12
13

14 Discussion

15
16 We investigated the impact of a number of variables related to L2 learning on proficient
17
18 L1 Spanish – L2 English bilinguals' brain responses to violations of English subject-verb
19
20 agreement. Although grand-mean analyses showed a reliable biphasic N400-P600 pattern,
21
22 individual analyses showed that this pattern was not representative of most learners' brain
23
24 responses. Individuals' brain responses varied along a continuum between N400- and P600-
25
26 dominant effects, with most participants showing dominance in one response or the other and
27
28 some showing relatively equal dominance. Thus, contrary to some models of L2 processing, a
29
30 large amount of variability in brain responses existed among the bilinguals we tested, despite
31
32 relatively high L2 proficiency and long-term L2 exposure. Importantly, multivariate analyses
33
34 showed that a large portion of this between-subject variability was systematic. By differentially
35
36 quantifying the type and magnitude of brain responses, we showed that different aspects of the
37
38 ERP signal (e.g., amplitude and polarity) were associated with different aspects of a learner's
39
40 background. Earlier age of arrival and higher motivation to speak like a native English speaker
41
42 were associated with greater dominance of the P600 effect, relative to the N400 effect.
43
44 Additionally, the linear combination of age of arrival, length of residence, frequency of use, and
45
46 motivation to speak in a native-like way accounted for a majority of the overall variance in the
47
48 type of brain response individuals' showed to agreement violations. However, contrary to
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 predictions made by the model proposed by Steinhauer and colleagues (Steinhauer et al., 2009),
4
5 proficiency had little impact on relative response dominance, as measured by the RDI, nor on
6
7 individuals' N400 or P600 magnitudes specifically. Instead, higher performance on the pen-and-
8
9 paper proficiency test was instead associated with a larger overall brain response across both
10
11 time windows, regardless of the type of response. It is important to note that we found these
12
13 partial correlations and multivariate effects despite a relatively small sample size. This indicates
14
15 exceptionally strong and systematic relationships, but also introduces a cautionary note in
16
17 interpreting the null effects seen in the regression analyses.
18
19
20
21

22 Our findings add to a growing number of L2 studies reporting N400 responses to
23
24 syntactic violations and extend these findings in important ways. Several longitudinal studies
25
26 have reported that syntactic anomalies elicit N400 effects in early-stage L2 learners, but elicit
27
28 P600 effects after increased L2 instruction (McLaughlin et al., 2010; Morgan-Short et al., 2010;
29
30 Morgan-Short, Steinhauer, Sanz, & Ullman, 2012; Osterhout et al., 2006). Furthermore, Tanner
31
32 and colleagues (Tanner et al., 2013) report cross-sectional results from novice L2 learners who
33
34 showed a nearly identical distribution of brain responses to that reported here. In most of these
35
36 studies, N400 effects are associated with the earliest stages of learning and are replaced by P600
37
38 effects within a single year of classroom L2 instruction. However, in the present study, N400
39
40 effects to L2 agreement violations persisted in some of the bilinguals, despite long-term L2
41
42 immersion and high L2 proficiency. There is therefore a seeming inconsistency between the
43
44 current results from immersed, proficient bilinguals and reports of low- and intermediate-
45
46 proficiency learners showing robust P600 (and no N400) effects to inflectional anomalies (Bond
47
48 et al., 2011; Frenck-Mestre et al., 2008; McLaughlin et al., 2010; Rossi et al., 2006; Tokowicz &
49
50 MacWhinney, 2005). Since N400 effects to syntactic violations have been associated with the
51
52
53
54
55
56
57
58
59
60

1
2
3 earliest stage of grammatical learning, it is conceivable that processing systems in the N400-
4
5 dominant learners in the current study may have somehow fossilized at an earlier stage of
6
7 acquisition. One argument might therefore be that N400 effects in late-stage bilinguals represent
8
9 instances of poor L2 learning, at least at a neurocognitive level.
10
11

12
13 The significant positive partial correlations between the response dominance index and
14
15 learners' age of arrival and motivation speak like a native speaker are consistent with this
16
17 hypothesis, as behavioral research has long noted relationships between these variables and long-
18
19 term learning outcomes. For example, Birdsong & Molis (2001) showed that among L1 Spanish
20
21 speakers of L2 English – a learner population similar to that tested here – increasing age of
22
23 arrival in an English-speaking environment was negatively correlated with performance on a
24
25 grammaticality judgment task, even among late learners who arrived after a putative critical
26
27 period for language learning (see also Flege et al., 1999; cf. Johnson & Newport, 1989). A
28
29 similar effect was found among our late learners; however, instead of showing a relationship
30
31 with behavioral measures of L2 knowledge, age of arrival correlated with the quality of brain
32
33 responses to syntactic anomalies in our ERP data (i.e., N400 or P600-dominance). Additionally,
34
35 our results provide evidence that motivation can impact learners' on-line brain responses during
36
37 language comprehension in a similar way, where P600-dominance correlated with higher
38
39 motivation to speak like a native speaker. Behavioral research has shown that learner motivation
40
41 correlates with increased L2 achievement (Gardner et al., 1997), even after controlling for
42
43 confounding factors such as age of acquisition and length of residence (Flege et al., 1999).
44
45 However, our measure did not differentiate between sub-types of learner motivation or provide a
46
47 detailed quantification of learner identity (see e.g., Dörnyei, 2005, for a thorough overview of
48
49 these issues), and future research may wish to more systematically investigate the role of
50
51
52
53
54
55
56
57
58
59
60

1
2
3 motivation on the neurocognition of an L2 (see Pulvermüller & Schumann, 1994, for one such
4
5 account). Nonetheless, our data show that higher motivation to speak like a native speaker, along
6
7
8 with an earlier age of arrival in an L2 environment, are associated with ERP responses which
9
10 more closely approximate those typically seen in grand mean waveforms in L1 experiments
11
12 investigating similar phenomena (e.g., Hagoort, Brown, & Groothusen, 1993; Hagoort & Brown,
13
14 2000; Osterhout & Mobley, 1995; Rossi, Gugler, Hahne, & Friederici, 2005).

15
16
17 However, a broader view of the relevant literature raises doubts about the “fossilization”
18
19 interpretation of the N400 effect in our bilinguals. Syntactic violations have sometimes elicited
20
21 N400 or biphasic N400-P600 responses similar to those reported here in L1 sentence processing
22
23 studies (e.g., Deutsch & Bentin, 2001; Mancini, Molinaro, Rizzi, & Carreiras, 2011a, 2011b;
24
25 Osterhout, 1997; Severens, Jansma, & Hartsuiker, 2008). As was the case in the current study,
26
27 some L1 studies have reported biphasic responses which were shown to be artifacts resulting
28
29 from averaging over individuals with different response profiles (Inoue & Osterhout, 2013;
30
31 Nieuwland & Van Berkum, 2008; Osterhout, 1997; Osterhout et al., 2004; see also Kos, van den
32
33 Brink, & Hagoort, 2012; Nakano, et al., 2010). Moreover, the “response dominance continuum”
34
35 reported here in bilinguals has also been observed in studies of L1 syntactic processing involving
36
37 Japanese-speaking adults (Inoue & Osterhout, 2013; see also Osterhout et al., 2004), as well as
38
39 English-speaking adults processing subject-verb agreement violations similar to those studied
40
41 here. For example, Tanner and Van Hell (2012) present findings from monolingual native
42
43 English speakers processing morphosyntactic violations similar to those used here. Although
44
45 grand mean waveforms in that study showed only a P600 effect to agreement anomalies,
46
47 subsequent analysis showed a similar continuum of brain responses to that reported here:
48
49 although most individuals showed a P600 effect to English morphosyntactic violations, nearly
50
51
52
53
54
55
56
57
58
59
60

1
2
3 one-third of participants showed an N400-dominant response. However, the early onset of the
4
5 P600 effect in some individuals cancelled out the N400 seen in others, resulting in little
6
7
8 negativity in the grand mean. Collectively, these results indicate that the individual differences
9
10 reported here exist in early-stage L2 learners, later-stage bilinguals, as well native speakers. That
11
12 is, individual differences in the N400-P600 continuum are not restricted to L2 learners, but
13
14 instead seem to be a property of language processing in general. Although age of arrival and
15
16 motivation were implicated in predicting brain response type in the L2 learners studied here,
17
18 these must be two of many interacting variables associated ERP response profiles across the
19
20 spectrum of language users.
21
22
23

24
25 Individual differences in adult language processing have previously been linked to
26
27 variation in working memory capacity, inhibitory control, neural efficiency, and cortical
28
29 connectivity, among other variables (e.g., Bates, Devescovi, & Wulfeck, 2001; Gernsbacher,
30
31 1993; Just & Carpenter, 1992; Van Petten, Weckerly, McIsaac, & Kutas, 1997; Prat, Keller, &
32
33 Just, 2007; Prat, 2011; Vos et al., 2001). Although these and other possibilities should be
34
35 explored further, the sentence processing model proposed by Ferreira, Christianson and
36
37 colleagues (Christianson, Hollingworth, Halliwell, & Ferreira, 2001; Ferreira, 2003) might prove
38
39 to be particularly efficacious for explaining the results reported here. Ferreira, Christianson and
40
41 colleagues argue that, in certain cases, comprehenders fail to compute detailed syntactic
42
43 representations of incoming linguistic information. In these circumstances individuals may
44
45 instead rely on shallower, lexically- or plausibility-based ‘good enough’ processing heuristics.
46
47 For example, in complex syntactic configurations such as passives or garden path sentences,
48
49 individuals may place greater weight on lexical or semantic than syntactic cues, especially when
50
51 the two sets of cues conflict (Ferreira, 2003; see also Kim & Osterhout, 2005). Seen in this
52
53
54
55
56
57
58
59
60

1
2
3 context, the processing of linguistic anomalies could proceed along two possible paths, one
4 based on lexical associations stored in memory (i.e., a shallower, ‘good enough’ analysis) or one
5 based on combinatorial, syntactic analyses. Relative reliance on one stream or the other might
6 then depend on the relative strength of lexical and syntactic cues in the sentence. The N400 and
7 P600 effects may therefore be seen as indexing the preferential engagement of one stream or the
8 other (Kuperberg, 2007; Osterhout et al., 2012; see Jackendoff, 2007; MacWhinney, Bates, &
9 Kliegl, 1984, for processing models which similarly propose a competitive dynamic between
10 lexical/conceptual and syntactic cues; see also Morgan-Short, Faretta-Stutenberg, Brill,
11 Carpenter, & Wong, in press; Ullman, 2004, 2005, for memory-based models which propose that
12 separable declarative and procedural memory systems may subserve these two processing
13 streams and which have implications for L2 learning).

14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29 Our results suggest further that cue strength and reliance on ‘good enough’ processing
30 heuristics may differ across individuals. Some individuals seem to rely more heavily on memory-
31 based heuristics (i.e., they are N400-dominant) and others on combinatorial information during
32 sentence processing (i.e., they are P600-dominant). Under this interpretation, the RDI represents
33 a metric of cue strength, a value that can vary across individuals at all levels of language
34 proficiency. For morphosyntactic dependencies like agreement, the relative weighting of the cues
35 (i.e., response dominance) seems to be especially malleable in many early-stage L2 learners,
36 such that additional instruction can produce a within-learner shift from an N400 response to a
37 P600 response to the same syntactic anomalies (McLaughlin et al., 2010). Presumably this
38 reflects the gradual acquisition of a grammatical rule and subsequent reduced reliance on lexical
39 information for dealing with that particular aspect of the L2. In L1 adults and later-stage L2
40 learners, response dominance may reflect more stable processing biases, though this
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 interpretation has yet to be verified empirically.
4

5 From this perspective, between-subject variability in the response-dominance continuum
6 for both L1 and L2 processing would be minimized by narrowing the set of relevant cues and
7 maximized by broadening the set of cues. For example, the syntactic violations in the current
8 study were realized with suppletive lexical alternations between short, high-frequency verbs (e.g.,
9 *is/are, has/have*). Those showing N400-dominant effects may have relied upon lexical form-
10 based expectations of verbal agreement features, while those showing P600-dominant responses
11 may have been more sensitive to the syntactic features carried on the verbs. The availability of
12 multiple cues to subject-verb agreement violations (wordform- versus syntactic feature-based)
13 may allow for variability in the exact cue that individuals attend to. Conversely, in contexts
14 where syntactic relationships are marked morphologically, the morphosyntactic cue to the
15 agreement dependency would be strengthened and violations should be more likely to elicit P600
16 effects across all individuals. One potential test of this hypothesis is provided by Foucart and
17 Frenc-Mestre (2012) who investigated grammatical gender processing in L1 English learners of
18 L2 French. In L2 learners, gender agreement violations elicited P600 effects when the anomaly
19 was detectable on adjectives, which are morphologically marked for gender, but an N400 effect
20 when the anomaly was detectable on nouns, where gender is encoded lexically and not
21 morphologically. The relative strength of phonological cues to morphosyntactic well-formedness
22 may also play a role in determining the neural signatures of processing agreement morphology.
23 Frenc-Mestre and colleagues (Frenc-Mestre et al., 2008) showed in L1 German learners of L2
24 French that visually-presented subject-verb agreement violations elicited a robust P600 effect
25 when there was a strong phonological cue to the violation, but a trend toward an N400 effect
26 when there was no phonological cue. Thus, a coalition of morphosyntactic and
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 morphophonological cues provided by target stimulus may play a role in preferentially engaging
4
5 one stream or the other.
6

7
8 Indeed, this may explain the discrepancy between the current findings and other L2
9
10 studies which have found P600 (but no N400) effects to morphosyntactic anomalies in learners
11
12 across the proficiency spectrum. P600 effects have been found most strongly in studies which
13
14 used agreement relationships marked by phonologically-realized, decomposable inflectional
15
16 morphemes (e.g., Frenck-Mestre et al., 2008; McLaughlin et al., 2010; Rossi et al., 2006;
17
18 Sabourin & Stowe, 2008; Tokowicz & MacWhinney, 2005; see also Hahne, Mueller, & Clahsen,
19
20 2006). Taken in this context, our findings suggest that later and less motivated learners may
21
22 prioritize lexical cues to syntactic relationships when they are available. However, given that
23
24 P600 effects are readily found in L2 learners across the proficiency spectrum, the broader L2
25
26 processing literature suggests that late learners are in no way restricted to these shallow
27
28 processes. Instead, our findings are consistent with the notion that individuals may modulate
29
30 their relative reliance on the lexical/semantic or morphosyntactic processing streams based on
31
32 the cues provided by the target stimulus.
33
34
35
36
37

38
39 Studies of L1 processing further corroborate this proposal. For example, Osterhout
40
41 (1997) showed that syntactically anomalous function words (which play a grammatical role but
42
43 have little referential meaning) elicited P600 effects in most participants. By contrast,
44
45 syntactically anomalous content words (which have rich referential meanings as well as
46
47 grammatical roles) elicited robust P600 effects in some subjects and N400 effects in others
48
49 (Osterhout, 1997). Similarly, anomalous Japanese case markers, which simultaneously indicate
50
51 the thematic and grammatical roles of each noun, elicit robust P600 effects in some Japanese-
52
53 speaking adults and N400 effects in others (Inoue & Osterhout, 2013; Osterhout et al., 2004). An
54
55
56
57
58
59
60

1
2
3 important implication of all of these results is that the grand-mean brain response observed in
4
5 any given experiment will be a function of not only the manipulated properties of the stimulus
6
7 (e.g., its linguistic category, morphological complexity, well-formedness) but also systematic
8
9 properties of the participants (e.g., each subject's weighting of the relevant linguistic cues), and
10
11 an interaction between the two. Repeated sampling from a population would by chance produce
12
13 different outcomes. Collectively, these results illustrate the dangers inherent in the exclusive use
14
15 of grand-average ERPs to characterize sentence processing. In some cases (as in the present
16
17 study), a thorough investigation of between-subject variability can provide an important
18
19 complement to traditional grand mean analyses.
20
21
22
23

24
25 In all of these experiments (including the present study), we have observed a robust
26
27 negative correlation between N400 and P600 magnitudes across individuals: as the magnitude of
28
29 one effect increased in size, the magnitude of the other effect decreased to a similar extent (Fig.
30
31 2). We have observed the negative correlation in L2 learners and in native speakers (Inoue &
32
33 Osterhout, 2013; Tanner, 2011; Tanner et al., 2013; Tanner & Van Hell, 2012; see also Nakano
34
35 et al, 2010; Nieuwland & Van Berkum, 2008; Osterhout 1997; Zhang, Li, Piao, Liu, Huang, &
36
37 Shu, 2013). Conceivably, the negative correlation manifests a dynamic interaction between
38
39 negatively correlated cortical networks involved in different aspects of sentence comprehension
40
41 (for example, lexico-semantic and grammatical aspects). Evidence to support this notion has
42
43 been acquired recently in an experiment that combined an ERP linguistic anomaly paradigm with
44
45 distributed source analysis (Inoue & Osterhout, 2013). This experiment demonstrated a negative
46
47 correlation in activity within anterior and posterior regions of the left-hemisphere perisylvian
48
49 cortex when participants read well-formed sentences. Individual differences in the “balance of
50
51 activation” within these areas when reading well-formed sentences predicted individual
52
53
54
55
56
57
58
59
60

1
2
3 differences in the ERP response to anomalous versions of those sentences, providing a
4
5 neurobiological explanation for the individual differences.
6
7

8 A further implication of the present findings follows from the observation that our two
9
10 ERP indices, the Response Magnitude Index and the Response Dominance Index, were
11
12 uncorrelated ($r = .076$). That is, variation in the quantity of the ERP response to linguistic
13
14 anomalies was uncorrelated with the quality of the response. Moreover, variability in each of
15
16 these measures independently was systematically associated with different subject-level
17
18 covariates, suggesting that the RDI and RMI together provide a more complete account of how
19
20 individual difference measures impact processing than traditional ERP measures alone, such as
21
22 mean amplitude. However, it remains to be seen how some of the individual variables mentioned
23
24 above (e.g., working memory, neural efficiency, cortical connectivity, etc.) map onto the
25
26 neurocognitive correlates of language processing that we report here (see Bond et al., 2011, for a
27
28 first attempt to link individuals' specific language aptitude and non-verbal reasoning ability with
29
30 L2 ERP effects). A further matter is the relationship between L1 and L2 processing mechanisms
31
32 within individuals. Given the evidence reviewed above that individuals vary along the
33
34 N400/P600 continuum even in their L1, it may be the case that an individual's location along the
35
36 continuum when processing his or her L1 may account for additional variance in L2 brain
37
38 responses. That is, do bilinguals show similar ERP response biases in their L1 and L2, or are the
39
40 mechanisms separable by language? While our current data cannot speak to this issue, it remains
41
42 a fruitful area for future research.
43
44
45
46
47
48
49

50 Finally, our findings show that ERP studies of L2 learning can inform theory beyond
51
52 simply identifying whether L2 syntactic processing can become 'native-like', where native-like
53
54 is defined as the presence of some specific brain response to a linguistic anomaly. While
55
56
57
58
59
60

1
2
3 traditional grand mean analyses are useful and informative about the central tendency and
4
5 normative brain response seen in a given population, a broad view of the L1 ERP and
6
7 neuroimaging literature indicates that native-likeness can encompass a continuum of brain
8
9 activation and ERP response profiles (Mason & Just, 2007; Nakano et al., 2010; Nieuwland &
10
11 Van Berkum, 2008; Osterhout, 1997; Pakulak & Neville, 2010; Prat, Keller, et al., 2007; Prat,
12
13 Long, & Baynes, 2007; Prat, 2011; Tanner & Van Hell, 2012; Tanner, 2011). Taken in this
14
15 context, our finding that variation exists in proficient L2 learners is unsurprising. The measures
16
17 we introduced here (the RDI and RMI) provide a new and useful way of understanding this
18
19 variation, particularly with regard to the multiple dimensions of processing that ERP recordings
20
21 provide. Although the pen-and-paper proficiency measures correlated with participants' brain
22
23 response magnitudes, the combination of response magnitude with response dominance
24
25 measures provided a much richer characterization of language comprehension profiles than
26
27 either off-line test batteries or traditional ERP quantification metrics would have provided alone.
28
29 Although further research is needed to better understand these relationships, as well as the extent
30
31 to which these findings will generalize to other learner populations or linguistic manipulations,
32
33 the present study provides yet another indication of the unique sensitivities of event-related
34
35 potentials as tools for studying language learning.
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

References

- 1
2
3
4
5
6
7
8
9 Allen, M., Badecker, W., & Osterhout, L. (2003). Morphological analysis in sentence
10 processing: An ERP study. *Language and Cognitive Processes, 18*, 405–430.
11
12
13
14 Bates, E., Devescovi, A., & Wulfeck, B. (2001). Psycholinguistics: A cross-language
15 perspective. *Annual Review of Psychology, 52*, 369–396.
16
17
18 Bentin, S. (1987). Event-related potentials, semantic processes, and expectancy factors in word
19 recognition. *Brain and Language, 31*, 308–327.
20
21
22
23 Van Berkum, J. J. A., Hagoort, P., & Brown, C. M. (1999). Semantic integration in sentences
24 and discourse: Evidence from the N400. *Journal of Cognitive Neuroscience, 11*, 657–671.
25
26
27
28 Birdsong, D., & Molis, M. (2001). On the evidence for maturational constraints in second-
29 language acquisition. *Journal of Memory and Language, 44*, 235–249.
30
31
32
33 Bond, K., Gabriele, A., Fiorentino, R., & Alemán Bañón, J. (2011). Individual differences and
34 the role of the L1 in L2 processing: An ERP investigation. In J. Herschensohn & D. Tanner
35 (Eds.), *Proceedings of the 11th Generative Approaches to Second Language Acquisition*
36 *Conference (GASLA 2011)* (pp. 17–29). Somerville, MA: Cascadilla Proceedings Project.
37
38
39
40 Carroll, J. B. (1962). The prediction of success in intensive foreign language training. In R.
41 Glaser (Ed.), *Training research and education* (pp. 87–36). Pittsburgh: University of
42 Pittsburgh Press.
43
44
45
46
47
48
49 Chen, L., Shu, H., Liu, Y., Zhao, J., & Li, P. (2007). ERP signatures of subject-verb agreement
50 in L2 learning. *Bilingualism: Language and Cognition, 10*, 161–174.
51
52
53
54 Christianson, K., Hollingworth, A., Halliwell, J. F., & Ferreira, F. (2001). Thematic roles
55 assigned along the garden path linger. *Cognitive Psychology, 42*, 368–407.
56
57
58
59
60

- 1
2
3 Deutsch, A., & Bentin, S. (2001). Syntactic and semantic factors in processing gender agreement
4
5 in Hebrew: Evidence from ERPs and eye movements. *Journal of Memory and Language*,
6
7 45, 200–224.
8
9
- 10 Ditman, T., Holcomb, P. J., & Kuperberg, G. R. (2007). An investigation of concurrent ERP and
11
12 self-paced reading methodologies. *Psychophysiology*, 44, 927–935.
13
14
- 15 Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second*
16
17 *language acquisition*. Mahwah, NJ: Erlbaum.
18
19
- 20 Dörnyei, Z., & Skehan, P. (2003). Individual differences in second language learning. In C. J.
21
22 Doughty & M. H. Long (Eds.), *The handbook of second language acquisition*. Malden,
23
24 MA: Blackwell.
25
26
- 27 Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology*, 47,
28
29 164–203.
30
31
- 32 Flege, J. E., Yeni-Komshian, G. H., & Liu, S. (1999). Age constraints on second-language
33
34 acquisition. *Journal of Memory and Language*, 41, 78–104.
35
36
- 37 Foucart, A., & Frenck-Mestre, C. (2012). Can late L2 learners acquire new grammatical
38
39 features? Evidence from ERPs and eye-tracking. *Journal of Memory and Language*, 66,
40
41 226–248.
42
43
- 44 Frenck-Mestre, C., Foucart, A., Carrasco, H., & Herschensohn, J. (2009). Processing of
45
46 grammatical gender in French as a first and second language. In L. Roberts, D. Véronique,
47
48 A. Nilsson, & M. Tellier (Eds.), *Eurosla Yearbook 9* (pp. 76–106). Amsterdam: Benjamins.
49
50
- 51 Frenck-Mestre, C., Osterhout, L., McLaughlin, J., & Foucart, A. (2008). The effect of
52
53 phonological realization of inflectional morphology on verbal agreement in French:
54
55 Evidence from ERPs. *Acta Psychologica*, 128, 528–536.
56
57
58
59
60

- 1
2
3 Friederici, A. D., Hahne, A., & Mecklinger, A. (1996). Temporal structure of syntactic
4 processing: Early and late event-related potential effects. *Journal of Experimental*
5
6 *Psychology: Learning, Memory, and Cognition*, 22, 1219–1248.
7
8
9
10 Gardner, R. C., Tremblay, P. F., & Masgoret, A.-M. (1997). Towards a full model of second
11 language learning: An empirical investigation. *The Modern Language Journal*, 81, 344–
12
13 362.
14
15
16
17 Gernsbacher, M. A. (1993). Less skilled readers have less efficient suppression mechanisms.
18 *Psychological Science*, 4, 294–298.
19
20
21
22 Gillon Dowens, M., Guo, T., Guo, J., Barber, H., & Carreiras, M. (2011). Gender and number
23 processing in Chinese learners of Spanish – Evidence from Event Related Potentials.
24
25 *Neuropsychologia*, 49, 1651–1659.
26
27
28
29 Gillon Dowens, M., Vergara, M., Barber, H., & Carreiras, M. (2010). Morphosyntactic
30 processing in late second-language learners. *Journal of Cognitive Neuroscience*, 22, 1870–
31
32 1887.
33
34
35
36 Hagoort, P., & Brown, C. M. (2000). ERP effects of listening to speech compared to reading:
37 The P600/SPS to syntactic violations in spoken sentences and rapid serial visual
38
39 presentation. *Neuropsychologia*, 38, 1531–1549.
40
41
42
43 Hagoort, P., Brown, C. M., & Groothusen, J. (1993). The syntactic positive shift as an ERP
44 measure of syntactic processing. *Language and Cognitive Processes*, 8, 439–484.
45
46
47
48 Hagoort, P., Hald, L., Bastiaansen, M., & Petersson, K. M. (2004). Integration of word meaning
49 and world knowledge in language comprehension. *Science*, 304, 438–441.
50
51
52
53 Hahne, A. (2001). What's different in second-language processing? Evidence from event-related
54 brain potentials. *Journal of Psycholinguistic Research*, 30, 251–266.
55
56
57
58
59
60

- 1
2
3 Hahne, A., & Friederici, A. D. (1999). Electrophysiological evidence for two steps in syntactic
4 analysis: Early automatic and late controlled processes. *Journal of Cognitive Neuroscience*,
5 *11*, 194–205.
6
7
8
9
10 Hahne, A., Mueller, J. L., & Clahsen, H. (2006). Morphological processing in a second language:
11 Behavioral and event-related brain potential evidence for storage and decomposition.
12 *Journal of Cognitive Neuroscience*, *18*, 121–134.
13
14
15
16
17 Handy, T. C. (2005). *Event-related potentials: A methods handbook*. Cambridge: MIT Press.
18
19
20 Van Herten, M., Chwilla, D. J., & Kolk, H. H. J. (2006). When heuristics clash with parsing
21 routines: ERP evidence for conflict monitoring in sentence perception. *Journal of Cognitive*
22 *Neuroscience*, *18*, 1181–1197.
23
24
25
26
27 Hopp, H. (2006). Syntactic features and reanalysis in near-native processing. *Second Language*
28 *Research*, *22*, 369–397.
29
30
31
32 Hopp, H. (2010). Ultimate attainment in L2 inflection: Performance similarities between non-
33 native and native speakers. *Lingua*, *120*, 901–931.
34
35
36
37 Inoue, K. & Osterhout, L. (2013). Sentence comprehension as a neural seesaw. Ms., University
38 of Washington.
39
40
41 Jackendoff, R. (2007). A Parallel Architecture perspective on language processing. *Brain*
42 *Research*, *1146*, 2–22.
43
44
45
46 Jackson, C. N., & Van Hell, J. G. (2011). The effects of L2 proficiency level on the processing of
47 wh-questions among Dutch second language speakers of English. *International Review of*
48 *Applied Linguistics in Language Teaching*, *49*, 195–219.
49
50
51
52
53 Jasper, H. H. (1958). The ten–twenty system of the International Federation.
54
55 *Electroencephalography and Clinical Neurophysiology*, *10*, 371–375.
56
57
58
59
60

- 1
2
3 Johnson, J., & Newport, E. (1989). Critical period effects in second language learning: The
4 influence of maturational state on the acquisition of English as a second language. *Cognitive*
5 *Psychology*, 21, 60–99.
6
7
8
9
10 Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual
11 differences in working memory. *Psychological Review*, 99, 122–149.
12
13
14
15 Kaan, E. (2002). Investigating the effects of distance and number interference in processing
16 subject-verb dependencies: An ERP study. *Journal of Psycholinguistic Research*, 31, 165–
17 193.
18
19
20
21
22 Kaan, E., Harris, A., Gibson, E., & Holcomb, P. J. (2000). The P600 as an index of syntactic
23 integration difficulty. *Language and Cognitive Processes*, 15, 159–201.
24
25
26
27 Kaan, E., & Swaab, T. Y. (2003). Repair, revision and complexity in syntactic analysis: An
28 electrophysiological differentiation. *Journal of Cognitive Neuroscience*, 15, 98–110.
29
30
31
32 Kim, A., & Osterhout, L. (2005). The independence of combinatorial semantic processing:
33 Evidence from event-related potentials. *Journal of Memory and Language*, 52, 205–225.
34
35
36
37 King, J. W., & Kutas, M. (1995). Who did what and when? Using word- and clause-level ERPs
38 to monitor working memory usage in reading. *Journal of Cognitive Neuroscience*, 7, 376–
39 395.
40
41
42
43
44 Kos, M., Van den Brink, D., & Hagoort, P. (2012). Individual variation in the late positive
45 complex to semantic anomalies. *Frontiers in Psychology*, 3:318.
46
47
48 doi:10.3389/fpsyg.2012.00318
49
50
51 Kuperberg, G. R. (2007). Neural mechanisms of language comprehension: Challenges to syntax.
52 *Brain Research*, 1146, 23–49.
53
54
55
56
57
58
59
60

- 1
2
3 Kuperberg, G. R., Caplan, D., Sitnikova, Tatiana, Eddy, M., & Holcomb, P. J. (2006). Neural
4 correlates of processing syntactic, semantic, and thematic relationships in sentences.
5
6 *Language and Cognitive Processes*, 21, 489–530.
7
8
9
10 Kuperberg, G. R., Holcomb, P. J., Sitnikova, T, Greve, D., Dale, A. M., & Caplan, D. (2003).
11 Distinct patterns of neural modulation during the processing of conceptual and syntactic
12 anomalies. *Journal of Cognitive Neuroscience*, 15, 272–293.
13
14
15
16
17 Kuperberg, G. R., Kreher, D. A., Sitnikova, Tatiana, Caplan, D., & Holcomb, P. J. (2007). The
18 role of animacy and thematic relationships in processing active English sentences: Evidence
19 from event-related potentials. *Brain and Language*, 100, 223–237.
20
21
22
23
24 Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in
25 language comprehension. *Trends in Cognitive Sciences*, 4, 463–470.
26
27
28
29 Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect
30 semantic anomaly. *Science*, 207, 203–205.
31
32
33
34 Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and
35 semantic association. *Nature*, 307, 161–163.
36
37
38
39 Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: (de)constructing
40 the N400. *Nature Reviews Neuroscience*, 9, 920–933.
41
42
43
44 Luck, S. J. (2005). *An introduction to the event-related potential technique*. Cambridge: MIT
45 Press.
46
47
48
49 MacCallum, R. C., Zhang, S., Preacher, K. J., & Rucker, D. D. (2002). On the practice of
50 dichotomization of quantitative variables. *Psychological Methods*, 7, 19–40.
51
52
53
54
55
56
57
58
59
60

- 1
2
3 MacWhinney, B., Bates, E., & Kliegl, R. (1984). Cue validity and sentence interpretation in
4
5 English, German, and Italian. *Journal of Verbal Learning and Verbal Behavior*, 23, 127–
6
7 150.
8
9
- 10 Mancini, S., Molinaro, N., Rizzi, L., & Carreiras, M. (2011a). A person is not a number:
11
12 Discourse involvement in subject-verb agreement computation. *Brain Research*, 1410, 64–
13
14 76.
15
16
- 17 Mancini, S., Molinaro, N., Rizzi, L., & Carreiras, M. (2011b). When persons disagree: An ERP
18
19 study of Unagreement in Spanish. *Psychophysiology*, 48, 1361–1371.
20
21
- 22 Mason, R. A., & Just, M. A. (2007). Lexical ambiguity in sentence comprehension. *Brain*
23
24 *Research*, 1146, 115–27.
25
26
- 27 McLaughlin, J., Tanner, D., Pitkänen, I., Frenck-Mestre, C., Inoue, K., Valentine, G., &
28
29 Osterhout, L. (2010). Brain potentials reveal discrete stages of L2 grammatical learning.
30
31 *Language Learning*, 60(Suppl. 2), 123–150.
32
33
- 34 Van de Meerendonk, N., Kolk, H. H. J., Vissers, C. T. W. M., & Chwilla, D. J. (2010).
35
36 Monitoring in language perception: Mild and strong conflicts elicit different ERP patterns.
37
38 *Journal of Cognitive Neuroscience*, 22, 67–82.
39
40
- 41 Molinaro, N., Barber, H., & Carreiras, M. (2011). Grammatical agreement processing in reading:
42
43 ERP findings and future directions. *Cortex*, 47, 908–930.
44
45
- 46 Molinaro, N., Kim, A., Vespignani, F., & Job, R. (2008). Anaphoric agreement violation: An
47
48 ERP analysis of its interpretation. *Cognition*, 106, 963–974.
49
50
- 51 Moreno, E. M., & Kutas, M. (2005). Processing semantic anomalies in two languages: An
52
53 electrophysiological exploration in both languages of Spanish-English bilinguals. *Cognitive*
54
55 *Brain Research*, 22, 205–220.
56
57
58
59
60

1
2
3 Morgan-Short, K., Faretta-Stutenberg, M., Brill, K., Carpenter, H., & Wong, P. C. M. (in press).

4
5 Declarative and procedural memory as individual differences in second language
6
7 acquisition. *Bilingualism: Language and Cognition*.

8
9
10 Morgan-Short, K., Sanz, C., Steinhauer, K., & Ullman, M. T. (2010). Second language
11
12 acquisition of gender agreement in explicit and implicit training conditions: An event-
13
14 related potentials study. *Language Learning*, 60, 154–193.

15
16
17 Morgan-Short, K., Steinhauer, K., Sanz, C., & Ullman, M. T. (2012). Explicit and implicit
18
19 second language training differentially affect the achievement of native-like brain activation
20
21 patterns. *Journal of Cognitive Neuroscience*, 24, 933–947.

22
23
24 Naiman, N., Fröhlich, M., Stern, H. H., & Todesco, A. (1996). *The good language learner*.
25
26 Philadelphia: Multilingual Matters.

27
28
29 Nakano, H., Saron, C., & Swaab, T. Y. (2010). Speech and span: Working memory capacity
30
31 impacts the use of animacy but not of world knowledge during spoken sentence
32
33 comprehension. *Journal of Cognitive Neuroscience*, 22, 2886–2898.

34
35
36 Nevins, A., Dillon, B., Malhotra, S., & Phillips, C. (2007). The role of feature-number and
37
38 feature-type in processing Hindi verb agreement violations. *Brain Research*, 1164, 81–94.

39
40
41 Newman, A. J., Tremblay, A., Nichols, E. S., Neville, H. J., & Ullman, M. T. (2012). The
42
43 influence of language proficiency on lexical semantic processing in native and late learners
44
45 of English. *Journal of Cognitive Neuroscience*, 24, 1205–1223.

46
47
48 Nieuwland, M. S., & Van Berkum, J. J. A. (2008). The interplay between semantic and
49
50 referential aspects of anaphoric noun phrase resolution: Evidence from ERPs. *Brain and*
51
52 *Language*, 106, 119–1131.

53
54
55
56
57
58
59
60

- 1
2
3 Oines, L., Miyake, A., & Kim, A. (2012). Individual differences in verbal working memory
4 predict reanalysis vs. integration in syntax-semantics conflict scenarios. Poster presented at
5
6 the Architectures and Mechanisms for Language Processing Conference. Riva del Garda,
7
8 Italy.
9
10
11
12 Ojima, S., Matsuba-Kurita, H., Nakamura, N., Hoshino, T., & Hagiwara, H. (2011). Age and
13 amount of exposure to a foreign language during childhood: Behavioral and ERP data on
14
15 the semantic comprehension of spoken English by Japanese children. *Neuroscience*
16
17 *Research, 70*, 197–205.
18
19
20
21
22 Ojima, S., Nakata, H., & Kakigi, R. (2005). An ERP study of second language learning after
23
24 childhood: Effects of Proficiency. *Journal of Cognitive Neuroscience, 17*, 1212–1228.
25
26
27
28 Osterhout, L. (1997). On the brain response to syntactic anomalies: Manipulations of word
29
30 position and word class reveal individual differences. *Brain and Language, 59*, 494–522.
31
32
33
34 Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic
35
36 anomaly. *Journal of Memory and Language, 31*, 785–806.
37
38
39
40 Osterhout, L., Kim, A., & Kuperberg, G. R. (2012). The neurobiology of sentence
41
42 comprehension. In M. Spivey, M. Joannisse, & K. McCrae (Eds.), *The Cambridge*
43
44 *handbook of Psycholinguistics* (pp. 365-389). Cambridge: Cambridge University Press.
45
46
47
48 Osterhout, L., McLaughlin, J., Pitkänen, I., Frenck-Mestre, C., & Molinaro, N. (2006). Novice
49
50 learners, longitudinal designs, and event-related potentials: A means for exploring the
51
52 neurocognition of second language processing. *Language Learning, 56*(Suppl. 1), 199–230.
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

- 1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
- Osterhout, L., & Nicol, J. (1999). On the distinctiveness, independence, and time course of the brain responses to syntactic and semantic anomalies. *Language and Cognitive Processes*, *14*, 283–317.
- Paczynski, M., & Kuperberg, G. R. (2011). Electrophysiological evidence for use of the animacy hierarchy, but not thematic role assignment, during verb-argument processing. *Language and Cognitive Processes*, *26*, 1402–1456.
- Pakulak, E., & Neville, H. J. (2010). Proficiency differences in syntactic processing of monolingual native speakers indexed by event-related potentials. *Journal of Cognitive Neuroscience*, *22*, 2728–2744.
- Van Petten, C. K., Weckerly, J., McIsaac, H. K., & Kutas, M. (1997). Working memory capacity dissociates lexical and sentential context effects. *Psychological Science*, *8*, 238–242.
- Prat, C. S. (2011). The brain basis of individual differences in language comprehension abilities. *Language and Linguistics Compass*, *5*, 635–649.
- Prat, C. S., Keller, T. A., & Just, M. A. (2007). Individual differences in sentence comprehension: A functional Magnetic Resonance Imaging investigation of syntactic and lexical processing demands. *Journal of Cognitive Neuroscience*, *19*, 1950–1963.
- Prat, C. S., Long, D. L., & Baynes, K. (2007). The representation of discourse in the two hemispheres: An individual differences investigation. *Brain and language*, *100*, 283–294.
- Pulvermüller, F., & Schumann, J. H. (1994). Neurobiological mechanisms of language acquisition. *Language Learning*, *44*, 681–734.
- Rossi, S., Gugler, M. F., Friederici, A. D., & Hahne, A. (2006). The impact of proficiency on syntactic second-language processing of German and Italian: Evidence from event-related potentials. *Journal of Cognitive Neuroscience*, *18*, 2030–2048.

- 1
2
3 Rossi, S., Gugler, M. F., Hahne, A., & Friederici, A. D. (2005). When word category information
4 encounters morphosyntax: An ERP study. *Neuroscience Letters*, 384, 228–233.
5
6
7 Sabourin, L., & Haverkort, M. (2003). Neural substrates of representation and processing of a
8 second language. In R. van Hout, A. Hulk, F. Kuiken, & R. Towell (Eds.), *The lexicon-*
9 *syntax interface in second language acquisition* (pp. 175–195). Amsterdam/Philadelphia:
10 John Benjamins.
11
12
13 Sabourin, L., & Stowe, L. A. (2008). Second language processing: When are first and second
14 languages processed similarly? *Second Language Research*, 24, 397–430.
15
16
17 Severens, E., Jansma, B. M., & Hartsuiker, R. J. (2008). Morphophonological influences on the
18 comprehension of subject-verb agreement: An ERP study. *Brain Research*, 1228, 135–144.
19
20
21 Skehan, P. (1989). *Individual differences in second-language learning*. New York: Arnold.
22
23
24 Steinhauer, K., White, E. J., & Drury, J. E. (2009). Temporal dynamics of late second language
25 acquisition: Evidence from event-related brain potentials. *Second Language Research*, 25,
26 13–41.
27
28
29 Tanner, D. (2011). Agreement mechanisms in native and nonnative language processing:
30 Electrophysiological correlates of complexity and interference. Ph.D. dissertation,
31 University of Washington.
32
33
34 Tanner, D., & Van Hell, J. G. (2012). ERPs reveal individual differences in syntactic processing
35 strategies. Poster presented at the 2012 Psychonomics Society Meeting, Minneapolis, MN.
36
37
38
39 Tanner, D., McLaughlin, J., Herschensohn, J., & Osterhout, L. (2013). Individual differences
40 reveal stages of L2 grammatical acquisition: ERP evidence. *Bilingualism: Language and*
41 *Cognition*, 16, 367–382.
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

- 1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
- Tanner, D., Nicol, J., Herschensohn, J., & Osterhout, L. (2012). Electrophysiological markers of interference and structural facilitation in native and nonnative agreement processing. In A. K. Biller, E. Y. Chung, & A. E. Kimball (Eds.), *Proceedings of the 36th annual Boston University Conference on Language Development* (pp. 594–606). Somerville, MA: Cascadilla Press.
- Tokowicz, N., & MacWhinney, B. (2005). Implicit and explicit measures of sensitivity to violations in second language grammar - An event-related potential investigation. *Studies in Second Language Acquisition*, 27, 173–204.
- Ullman, M. T. (2001). The neural basis of lexicon and grammar in first and second language: The declarative/procedural model. *Bilingualism: Language and Cognition*, 4, 105–122.
- Ullman, M. T. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition*, 92, 231–270.
- Ullman, M. T. (2005). A cognitive neuroscience perspective on second language acquisition: The declarative/procedural model. In C. Sanz (Ed.), *Mind and context in adult second language acquisition* (pp. 141–178). Washington, DC: Georgetown University Press.
- Van Hell, J. G., & Tanner, D. (2012). Second language proficiency and cross-language lexical activation. *Language Learning*, 62(Suppl. 2), 148–171.
- Van Hell, J. G., & Tokowicz, N. (2010). Event-related brain potentials and second language learning: Syntactic processing in late L2 learners at different L2 proficiency levels. *Second Language Research*, 26, 43–74.
- Vos, S. H., Gunter, T. C., Kolk, H. H. J., & Mulder, G. (2001). Working memory constraints on syntactic processing: An electrophysiological investigation. *Psychophysiology*, 38, 41–63.

- 1
2
3 Weber-Fox, C. M., & Neville, H. J. (1996). Maturational constraints on functional
4 specializations for language processing: ERP and behavioral evidence in bilingual speakers.
5
6 *Journal of Cognitive Neuroscience*, 8, 231–256.
7
8
9
- 10 Weber-Fox, C. M., & Neville, H. J. (1999). Functional neural subsystems are differentially
11 affected by delays in second language immersion: ERP and behavioral evidence in
12
13 bilinguals. In D. Birdsong (Ed.), *Second language acquisition and the critical period*
14
15 *hypothesis*. Mahwah, NJ: Erlbaum.
16
17
18
- 19 Weckerly, J., & Kutas, M. (1999). An electrophysiological analysis of animacy effects in the
20 processing of object relative sentences. *Psychophysiology*, 36, 559–570.
21
22
23
- 24 Zhang, Y., Li, P., Piao, Q., Liu, Y., Huang, Y., & Shu, H. (2013). Syntax does not necessarily
25 precede semantics in sentence processing: ERP evidence from Chinese. *Brain and*
26
27 *Language*, 126, 8–19.
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Footnotes

¹ Some studies of morphosyntactic processing have reported an additional negative-going wave prominent over left-anterior portions of the scalp, with an onset around between 100-400ms poststimulus (the LAN effect: e.g., Friederici et al., 1996; Osterhout & Holcomb, 1992; Osterhout & Mobley, 1995; Rossi, Gugler, Hahne, & Friederici, 2005). However, the LAN has been inconsistent across studies (e.g., Allen, Badecker, & Osterhout, 2003; Ditman, Holcomb, & Kuperberg, 2007; Kaan, 2002; Kaan, Harris, Gibson, & Holcomb, 2000; Kaan & Swaab, 2003; Kuperberg et al., 2003; Molinaro, Kim, Vespignani, & Job, 2008; Nevins, Dillon, Malhotra, & Phillips, 2007; Osterhout & Mobley, 1995), and more research is needed to precisely identify the experimental conditions under which LANs are reliably elicited. We therefore focus on the P600 as an index of morphosyntactic processing.

²This time window begins somewhat later than the 300-500ms time window frequently used to quantify the N400 in studies of monolingual language processing. However, N400 effects have reported to be somewhat delayed in some studies of L2 processing, relative to L1 processing (Hahne, 2001).

1
2
3 Figure Captions
4
5

6
7 *Figure 1.* Grand mean waveforms for grammatical (black line) and ungrammatical (red line)
8 verbs from nine representative electrodes. Onset of the critical verb is indicated by the vertical
9 bar. Calibration bar shows 3 μ V of activity; each tick mark represents 100ms of time. Positive
10 voltage is plotted down.
11
12
13
14
15

16
17 *Figure 2.* Scatterplot showing the distribution of N400 and P600 effect magnitudes across
18 participants, averaged within a centro-parietal ROI (C3, Cz, C4, P3, Pz, P4). Each dot represents
19 a data point from a single participant. The solid line shows the best-fit line for the data from the
20 regression analysis. The dashed line represents equal N400 and P600 effect magnitudes:
21
22 individuals above/to the left of the dashed line showed primarily an N400 effect to agreement
23 violations, while individuals below/to the right of the dashed line showed primarily a P600 effect.
24
25
26
27
28
29
30
31
32
33

34 *Figure 3.* Grand mean waveforms and scalp topographies of grammaticality effects (i.e., mean
35 amplitude in ungrammatical minus grammatical condition) presented separately for those who
36 showed primarily N400 (Panel A) and P600 (Panel B) effects to subject-verb agreement
37 violations. Onset of the critical verb is indicated by the vertical bar. Calibration bar shows 3 μ V
38 of activity; each tick mark represents 100ms of time. Positive voltage is plotted down.
39
40
41
42
43
44
45
46 Topographic map scale is in microvolts.
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Table 1. *Background information for participants*

Measure	Mean	St. Dev.	Range
Age at testing (years)	35.2	6.9	24 – 49
AoE (years)	11.6	7.1	5 – 30
AoA (years)	23.9	6.4	15 – 40
LoR (years)	10.6	6.4	5 – 27
<i>Self-rated Proficiency (Spanish)</i>			
Speaking	7	0	7 – 7
Listening	7	0	7 – 7
Reading	7	0	7 – 7
Writing	6.95	0.2	6 – 7
<i>Self-rated Proficiency (English)</i>			
Speaking	5.8	0.7	4 – 7
Listening	6.2	0.7	4 – 7
Reading	6.3	0.6	5 – 7
Writing	5.8	0.9	4 – 7
<i>English ECPE Proficiency Scores</i>			
Vocabulary (30 possible)	27.7	2.7	20 – 30
Grammar (20 possible)	17.4	2.3	11 – 20
Total (50 possible)	45.1	4.2	36 – 50
L2 Motivation	6.4	0.8	5-7
Frequency use English (overall)	4.6	0.9	3-7

Table 2. *Mean, standard deviation, and range for DVs and background measures used in regression models*

Measure	Mean	St. Dev.	Range
RDI	0.25	3.28	-5.40 – 7.93
RMI	3.55	1.71	0.62 – 8.14
LogAoA	1.36	0.11	1.18 – 1.60
LogLoR	1.00	0.22	0.70 – 1.43
Freq.	4.60	0.88	3 – 7
Proficiency	45.05	4.17	36 – 50
Motivation	0.55	0.51	0 – 1

Note: RDI = Response Dominance Index; RMI = Response Magnitude Index; LogAoA = Log-Age of Arrival; LogLoR = Log-Length of Residence; Freq. = Frequency of L2 Use; Proficiency = Total ECPE Score; Motivation = Dichotomized Motivation self-report.

Table 3. *Correlation matrix for DVs and background measures used in regression models*

	RDI	RMI	LogAoA	LogLoR	Freq.	Proficiency	Motivation
RDI	--						
RMI	.076	--					
LogAoA	-.277	-.327	--				
LogLoR	.445*	.097	-.367	--			
Freq.	.286	.104	-.164	.235	--		
Proficiency	.007	.478*	-.003	.264	.077	--	
Motivation	.517*	.031	.385	.016	.047	.209	--

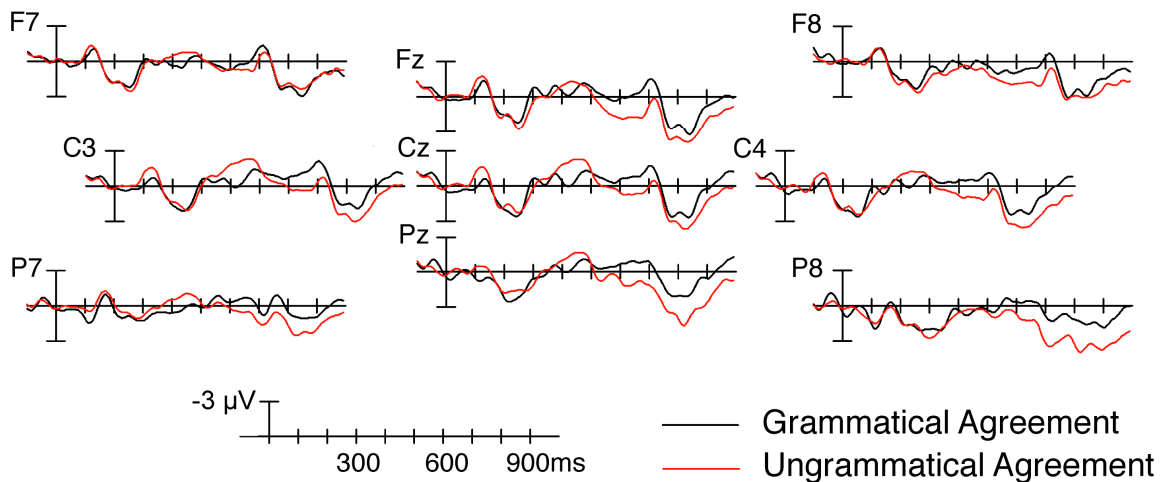
Note: RDI = Response Dominance Index; RMI = Response Magnitude Index; LogAoA = Log-Age of Arrival; LogLoR = Log-Length of Residence; Freq. = Frequency of L2 Use; Proficiency = Total ECPE Score; Motivation = Dichotomized Motivation self-report. ⁺ $p < .06$, * $p < .05$.

Table 4. Multiple regression coefficients for model with RDI as DV

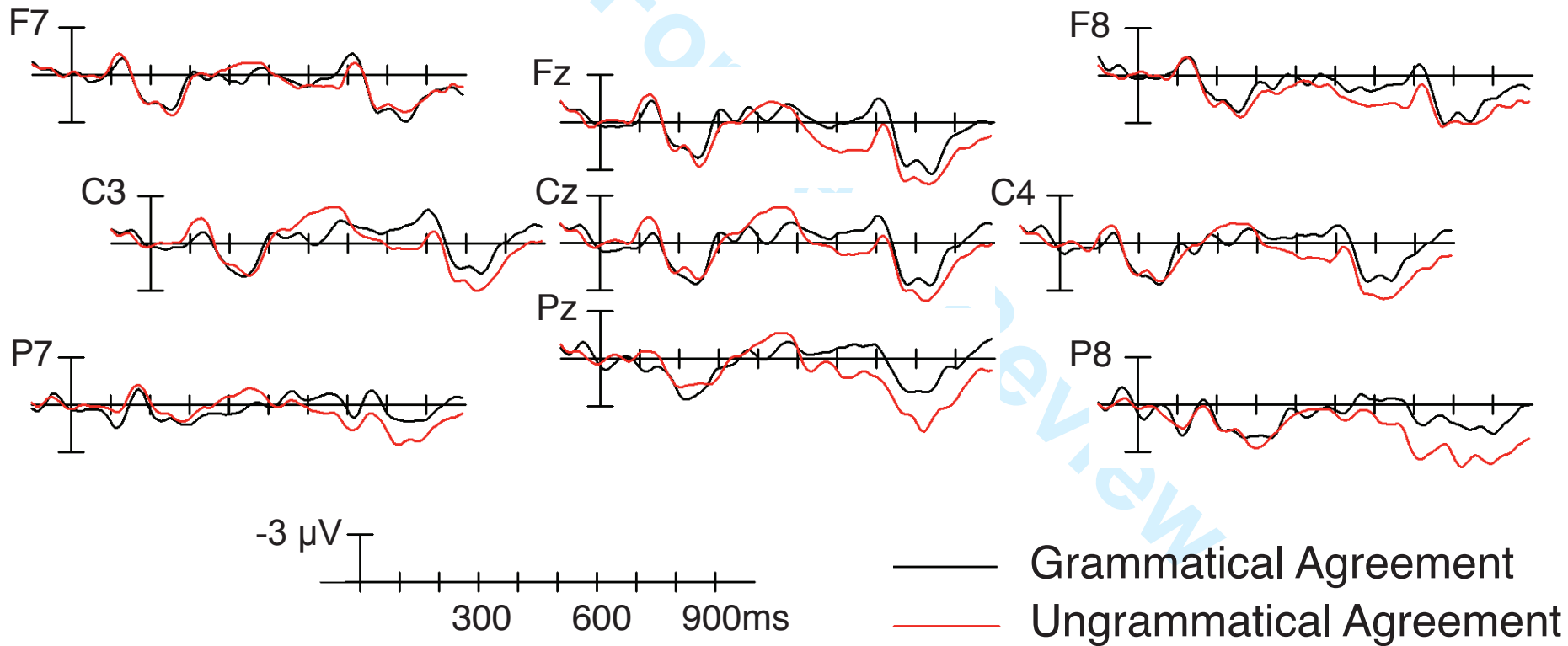
Predictor	B	SE	β	t	Partial-r	<i>p</i>
LogAoA	-13.20	5.43	-0.42	-2.25	-.515	.041
LogLoR	4.608	2.64	0.31	1.75	.423	.103
Freq.	.478	0.60	0.13	0.80	.208	.439
Proficiency	-.186	0.13	-0.24	-1.43	-.357	.175
Motivation	4.60	1.28	0.72	4.08	.737	.001

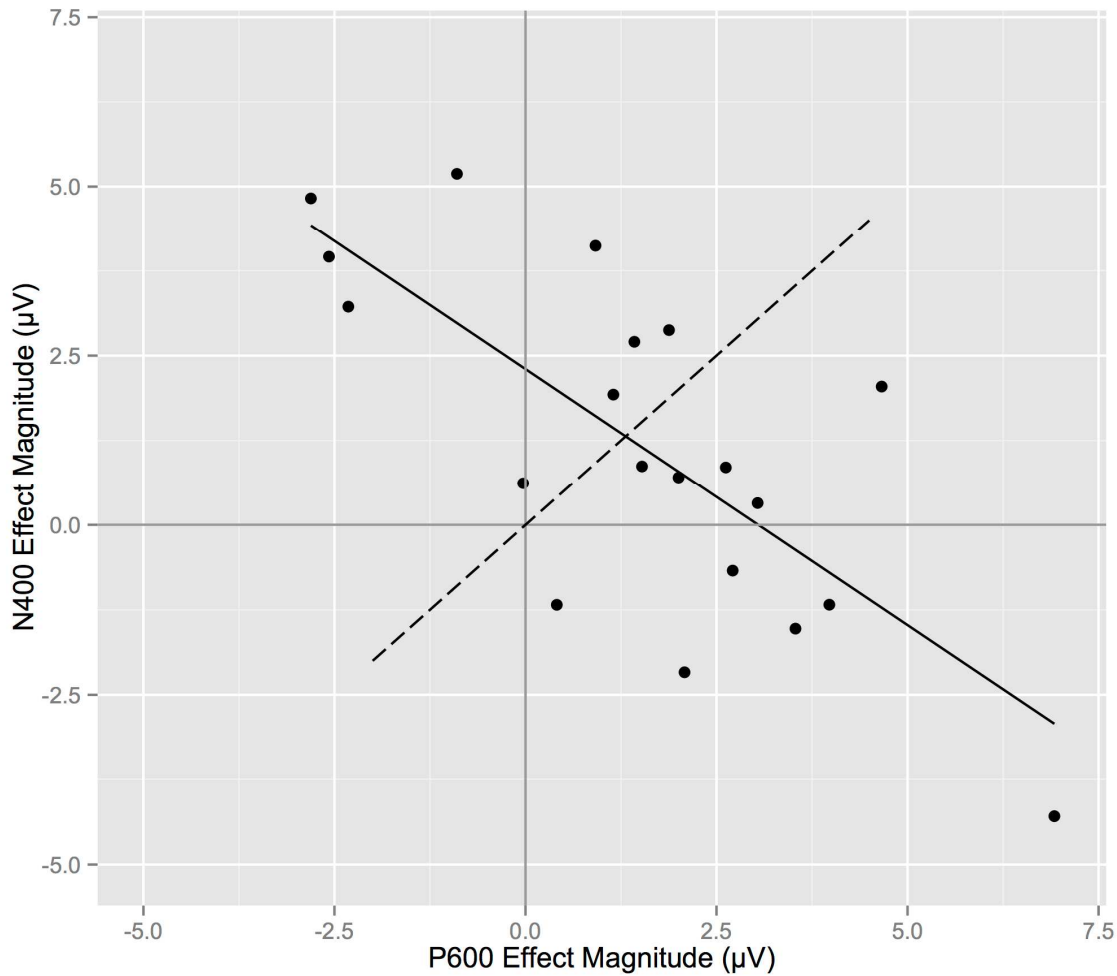
Note: $R^2 = .660$, Adjusted $R^2 = .539$, $F(5, 14) = 5.434$, $p = .006$. B and β represent the unstandardized and standardized regression coefficients, respectively.

Peer Review

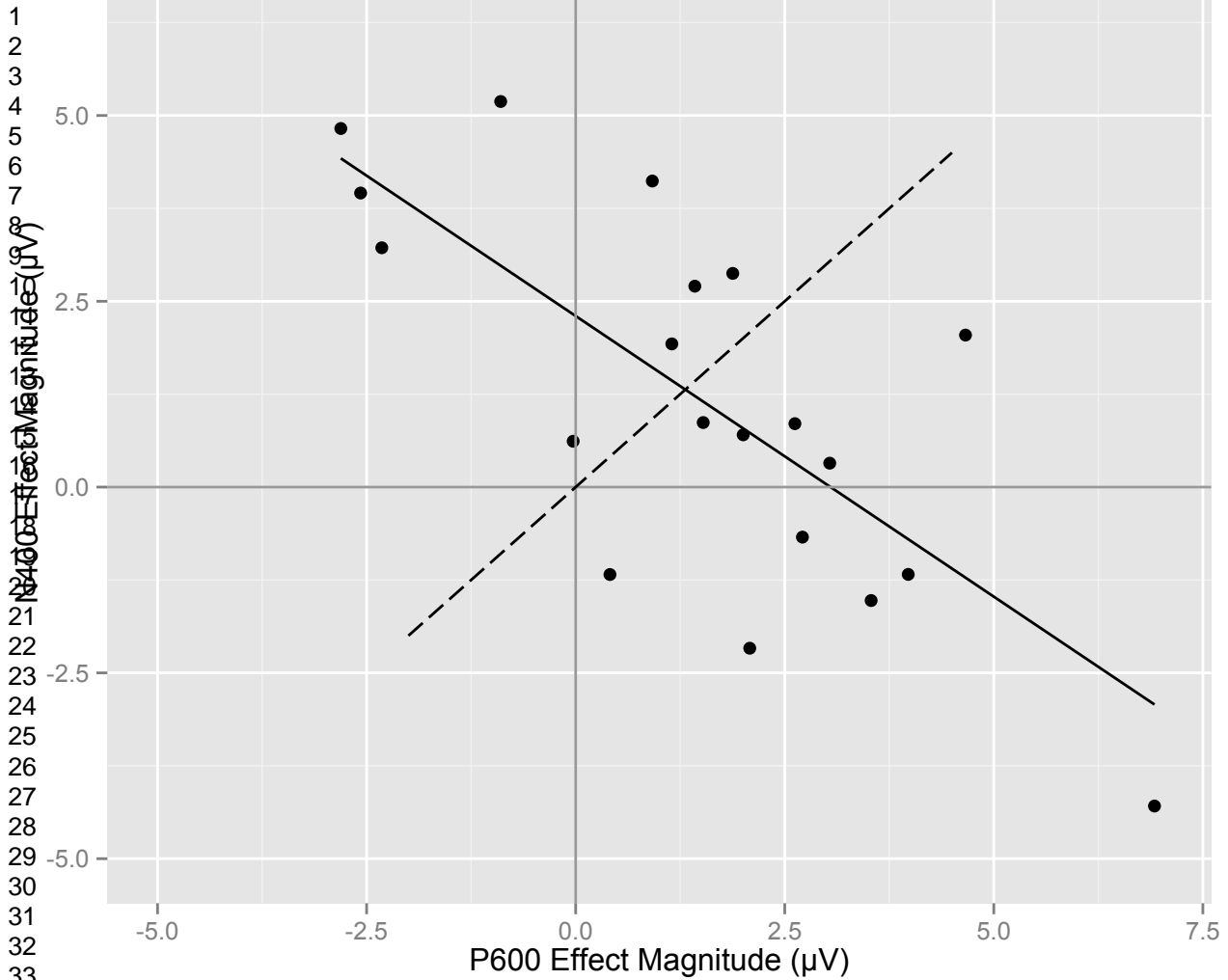


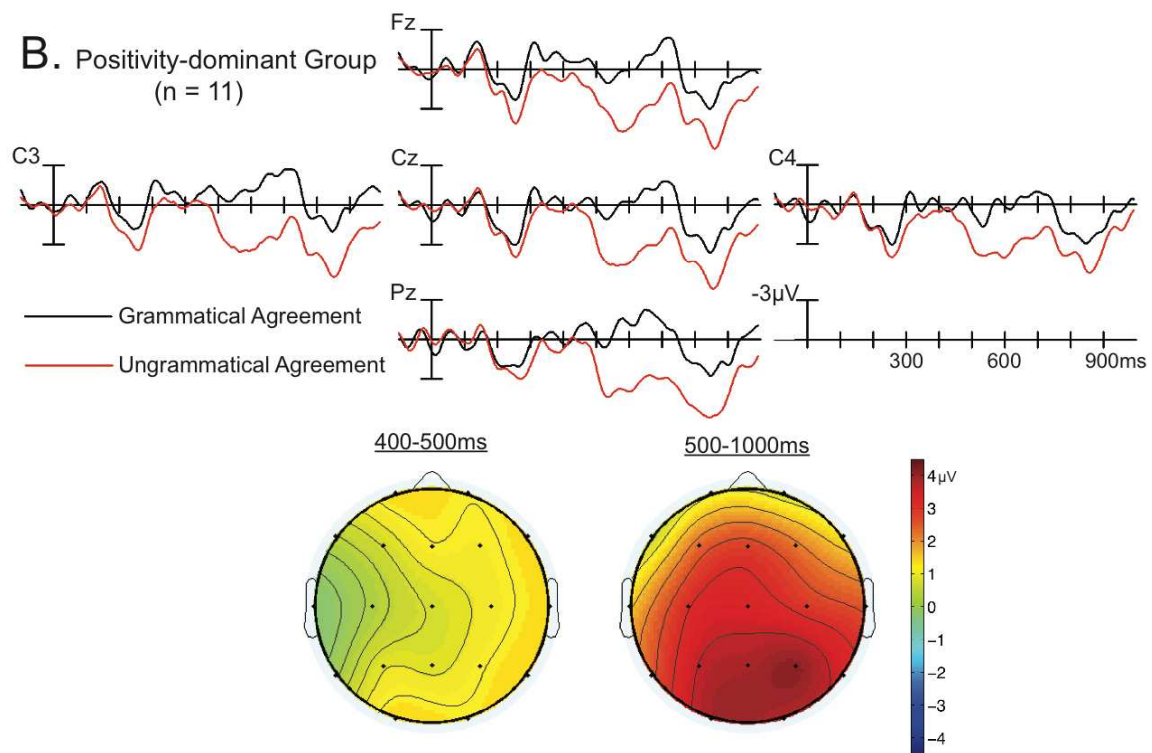
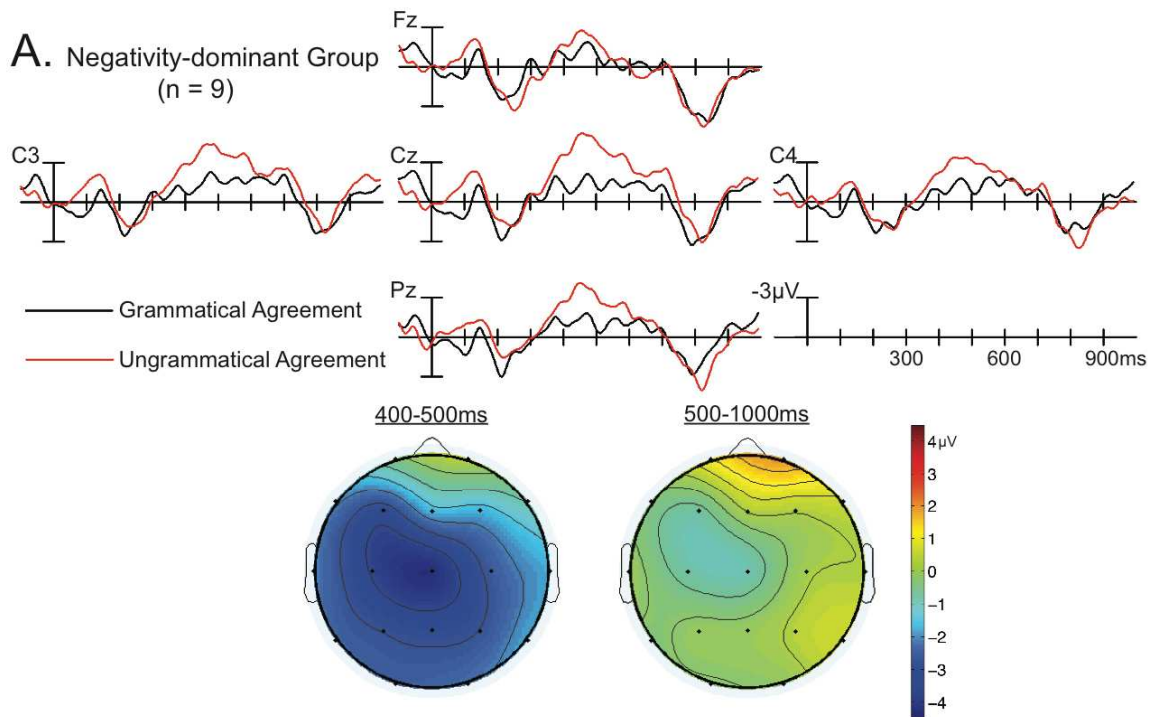
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60





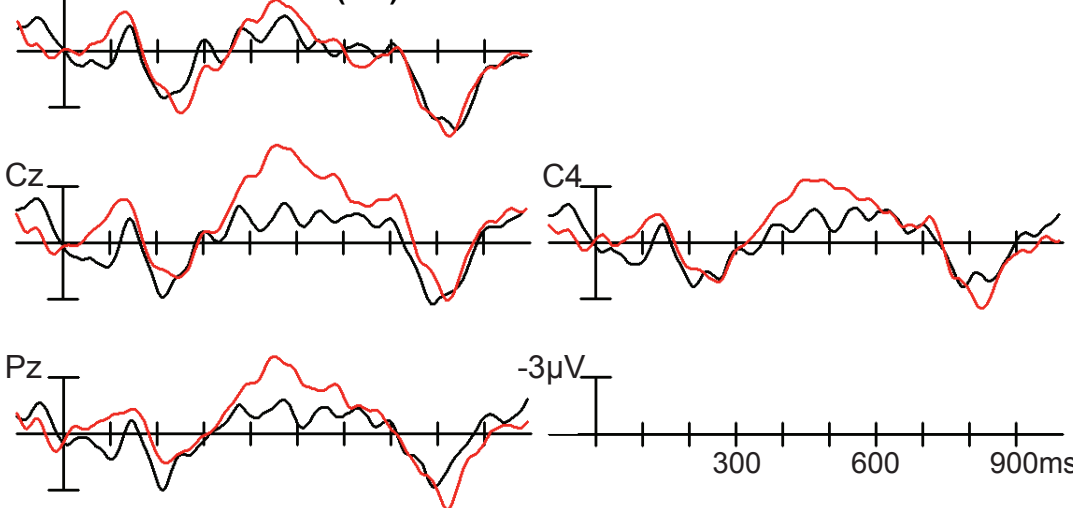
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60





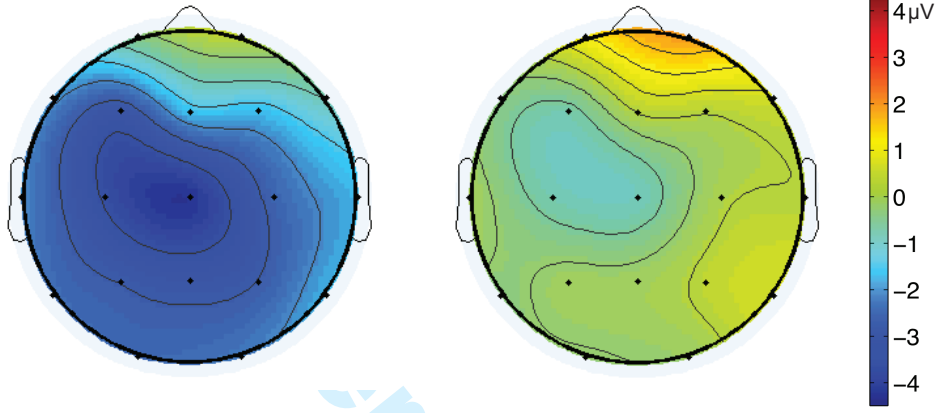
A. Negativity-dominant Group (n = 9)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31



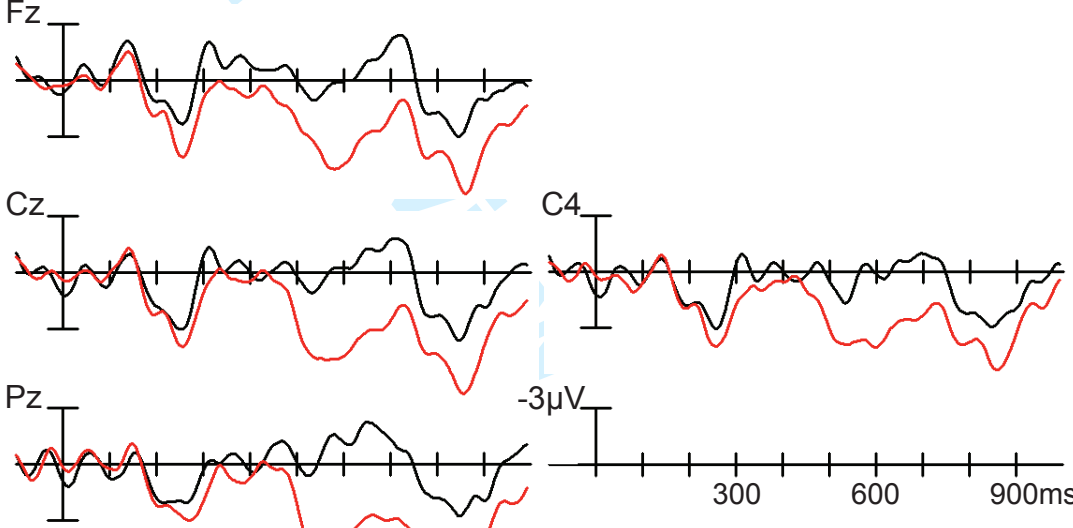
400-500ms

500-1000ms



B. Positivity-dominant Group (n = 11)

32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



400-500ms

500-1000ms

