Version of Record: https://www.sciencedirect.com/science/article/pii/S0028393216303803
Manuscript_61b7755d36b8f8a463d58a2e7eda8b2d

Native-language N400 and P600 predict dissociable language-learning abilities in adults

Zhenghan Qi*[a], Sara D. Beach*[b], Amy S. Finn[a,1] Jennifer Minas[a], Calvin Goetz[a], Brian Chan[a], and John D.E. Gabrieli[a,c]

[a] McGovern Institute for Brain Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[b] Program in Speech and Hearing Bioscience and Technology, Division of Medical Sciences, Harvard University, Cambridge, MA 02138, USA
[c] Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

[1] Present address: Department of Psychology, University of Toronto, Canada

* Authors contributed equally to this work.

Authors' information
Zhenghan Qi: zqi@mit.edu
Sara D. Beach: beach@g.harvard.edu
Amy S. Finn: amy.finn@utoronto.ca
Jennifer Minas: jminas@mit.edu
Calvin Goetz: cgoetz@mit.edu
Brian Chan: brianc07@mit.edu
John D. E. Gabrieli: gabrieli@mit.edu

Corresponding author
Zhenghan Qi
Massachusetts Institute of Technology
43 Vassar St., 46-4033
Cambridge, MA 02139, USA
Email: zqi@mit.edu
Tel: +1-617-324-4959
Fax: +1-617-324-5311

© 2017 published by Elsevier. This manuscript is made available under the Elsevier user license
https://www.elsevier.com/open-access/userlicense/1.0/

ABSTRACT

Language learning aptitude during adulthood varies markedly across individuals. An individual's native-language ability has been associated with success in learning a new language as an adult. However, little is known about *how* native-language processing affects learning success and what neural markers of native-language processing, if any, are related to success in learning. We therefore related variation in electrophysiology during native-language processing to success in learning a novel artificial language. Event-related potentials (ERPs) were recorded while native English speakers judged the acceptability of English sentences prior to learning an artificial language. There was a trend towards a double dissociation between native-language ERPs and their relationships to novel syntax and vocabulary learning. Individuals who exhibited a greater N400 effect when processing English semantics showed better future learning of the artificial language overall. The N400 effect was related to syntax learning via its specific relationship to vocabulary learning. In contrast, the P600 effect size when processing English syntax predicted future syntax learning but not vocabulary learning. These findings show that distinct neural signatures of native-language processing relate to dissociable abilities for learning novel semantic and syntactic information.

1. INTRODUCTION

Although typically developing children seem to acquire their native language effortlessly, learning a novel language in adulthood is notoriously challenging and marked by wide disparities in attainment (Newport, 1990; Snow & Hoefnagel-Hohle, 1978). Several decades of research have examined the cognitive/behavioral, affective/motivational, and instructional/environmental factors that may relate to better learning outcomes (Carroll et al., 2010; Linck et al., 2013; Melby-Lervåg & Lervåg, 2011; see also reviews: DeKeyser, 2012; Lesaux et al., 2006; Sparks, 2012). It is striking that, across numerous studies, native-language skills form the basis of language-learning aptitude (Carroll et al., 2010) and are strongly related to novel-language attainment (Melby-Lervåg & Lervåg, 2011; Sparks et al., 2012; Sparks et al., 2009). Even distinct subskills such as phonological processing and language comprehension (including semantics and syntax) appear to be shared across native-language and novel-language learning (Sparks et al., 2012). However, semantics and syntax are intertwined in the process of language acquisition as well as in many language assessments. It is not yet known how individuals' processing of native-language semantics and syntax impacts their success in learning those aspects of a novel language.

Numerous studies of native-language processing have revealed dissociable neural substrates of semantic and syntactic processing. Lesion studies dating from the work of Broca and Wernicke provide evidence for separate loci of semantic and syntactic processing, owing to the variety of aphasia that results from the insult (Alexander et al., 1990; Benson, 1985; Vignolo, 1988); neuropsychological approaches also support the claim that semantic and syntactic knowledge rely on separate memory systems (Ullman et al., 1997). Event-related potentials

(ERPs) in response to semantic and syntactic features exhibit distinct temporal and scalp distribution patterns (Hagoort et al., 1993; Kutas & Hillyard, 1984; Osterhout & Holcomb, 1992). Moreover, as individuals' semantic and syntactic skills improve during foreign-language learning, these ERPs appear to grow more native-like (Batterink & Neville, 2013; Hahne et al., 2006; McLaughlin et al., 2004; McLaughlin et al., 2010; Midgley et al., 2009; Morgan-Short et al., 2012a; Newman et al., 2012; Osterhout et al., 2006; Tanner et al., 2014; Yum et al., 2014).

Given that there are distinct neural signatures of semantic and syntactic processing, it is possible that these neural indices could be predictive of one's ability to learn these two aspects of a novel language. In this study, we measured ERPs in the native language and related them to adults' learning of a miniature artificial language in an immersive training setting.

1.1 ERP Markers of Native Language Processing

The N400 and P600 are among the most studied electrophysiological signatures of native language comprehension. The N400 is a negative deflection peaking around 400 ms after stimulus onset that indexes the processing of meaning. A greater N400 amplitude can be elicited by manipulating cloze probability (DeLong et al., 2005; Federmeier et al., 2007; Kutas & Hillyard, 1984; Kutas, 1993), semantic priming (Bentin et al., 1985; Kutas, 1993), word frequency (Van Petten & Kutas, 1990), semantic category (Federmeier & Kutas, 1999), and orthographic neighborhood size (Laszlo & Federmeier, 2009). In contrast, the P600 is a positive deflection that usually begins around 500 ms after stimulus onset and lasts a few hundred milliseconds. P600 amplitude reflects aspects of the processing of syntax and increases in response to various types of morphosyntactic violations (Hagoort et al., 1993; Osterhout & Holcomb, 1992), more complicated syntax (Friederici et al., 2002; Kaan et al., 2000), and less preferred syntactic structure (Itzhak et al., 2010; Osterhout et al., 1994).

Syntactic processing has also been associated with activity in an earlier time window, between 250 and 500 ms after stimulus onset. The widely-reported left anterior negativity (LAN) or anterior negativity (AN) is elicited by morphosyntactic violations such as subject-verb agreement errors and case marking errors (Friederici, Hahn, & Mecklinger, 1996; Friederici & Frisch, 2000; Coulson et al., 1998; Münte, Heinze & Mangun, 1993; Münte et al., 1997; Gunter, Stowe, & Mulder, 1997). The LAN/AN is often interpreted as reflecting neural sensitivity to morphosyntactic errors in an intermediate phase of sentence processing and, together with a later positivity (the P600), may contribute to a biphasic LAN-P600 pattern (Steinhauer & Drury, 2002). Additionally, an N400 elicited by morphosyntactic violations has been found to reflect processing at a lexical level (Osterhout, 1997; Tanner & Van Hell, 2014). Lastly, a positive deflection with a scalp distribution like that of the P600 in this same time window (the "early P600") has also been reported to reflect immediate diagnosis of and recovery from a nonpreferred structure (Mecklinger et al., 1995; Patel et al., 1998).

Interestingly, despite behavioral fluency, adults display substantial variability in electrophysiological responses to inputs in their native language. Some variability in each component has been associated with variation in specific language skills. A larger magnitude of the P600 has been related to higher grammatical proficiency in adults (Pakulak & Neville, 2010), and a greater N400 effect has been linked to comprehension skill (Landi & Perfetti, 2007). A few notable studies describe how individuals differentially engage semantic versus syntactic processes in response to the same stimuli. In sentences with temporary syntactic ambiguity (e.g., *The boat sailed down the river sank*), the disambiguating word *sank* elicited P600s in some participants and N400s in others (Osterhout, 1997). The seemingly qualitative individual difference was attributed to individuals' sensitivity to the semantic versus syntactic

consequences of the critical word (Osterhout, 1997). In a more strictly syntactic framework, a quantitative continuum of N400- to P600-like effects was reported in response to morphosyntactic violations (e.g., **The clerk were unhappy*; **The crime rate was increase*) (Tanner & Van Hell, 2014). The variability in participants' responses, like that reported by Osterhout (1997), suggests individual differences in processing schemes: based on word form (N400-types), combinatorial constraints (P600-types), or both (Tanner & Van Hell, 2014). A similar N400-P600 continuum has also been reported in morphosyntactic processing in a study of Japanese native speakers (Tanner et al., 2014). All three studies argue for the existence of individual differences in adults' native-language processing, the ramifications of which have not yet been made clear.

1.2 Predictors of Novel Language Learning Aptitude

Brain imaging has contributed much to our understanding of how the adult brain supports or constrains the learning of linguistic information. MRI studies have related pre-training anatomical features and functional neural activity (see review: Li & Grant, 2015) to individuals' ability to learn new phonetic contrasts (Golestani et al., 2007; Ventura-Campos et al., 2013), words (López-Barroso et al., 2013; Tan et al., 2011; Veroude et al., 2010; Wong et al., 2011; Wong et al., 2007), syntax (Flöel et al., 2009; Loui et al., 2011), and holistic language skills (Qi et al., 2015). These studies show that structural and functional characteristics of the brain prior to learning are associated with various learning outcomes. However, the relevant brain regions and white-matter tracts reported in these studies are mostly overlapping, and thus not specific to the subskills of language learning. Existing ERP studies of adult language learning mainly track emerging novel-language skills and have established separable ERP markers for semantic and syntactic learning outcomes (McLaughlin et al., 2004; McLaughlin et al., 2010; Mestres-Misse et

al., 2006; Tanner et al., 2014; White et al., 2012; Yum et al., 2014). Nevertheless, none of these studies address whether individuals are equipped with specific learning strengths for semantic and syntactic information before such learning begins.

To bridge the gap between the neural signatures of native language processing and novel-language aptitude, we measured individuals' neural sensitivity to native-language semantic and syntactic anomalies using scalp ERPs. We then immersed learners in a novel language, exposing them to new vocabulary and syntactic information simultaneously in a three-day naturalistic learning paradigm. We hypothesized that the neural index of native-language semantic processing (the N400) would be related to individuals' ability to learn vocabulary, which supports their ability to make accurate semantic acceptability judgments in the new language, and that the neural index of native-language syntactic processing (the P600) would be related to individuals' ability to learn the morphosyntax of a new language.

2. MATERIAL AND METHODS

2.1 Overview of the language training procedure

Participants completed a pre-training EEG session on Day 1 of the study, prior to training and testing on the miniature artificial language (Fig. 1). Participants then took part in the language training sessions on Day 1 (subsequent to the English EEG task), Day 2, and Day 3 of the experiment. Proficiency tests on vocabulary, semantic, and syntactic learning outcomes were administered immediately following each training session. Participants also returned on Day 4 to repeat the tests, although there was no training on that day. Together, the four-day procedure spanned no more than six calendar days. Details of each session follow below.

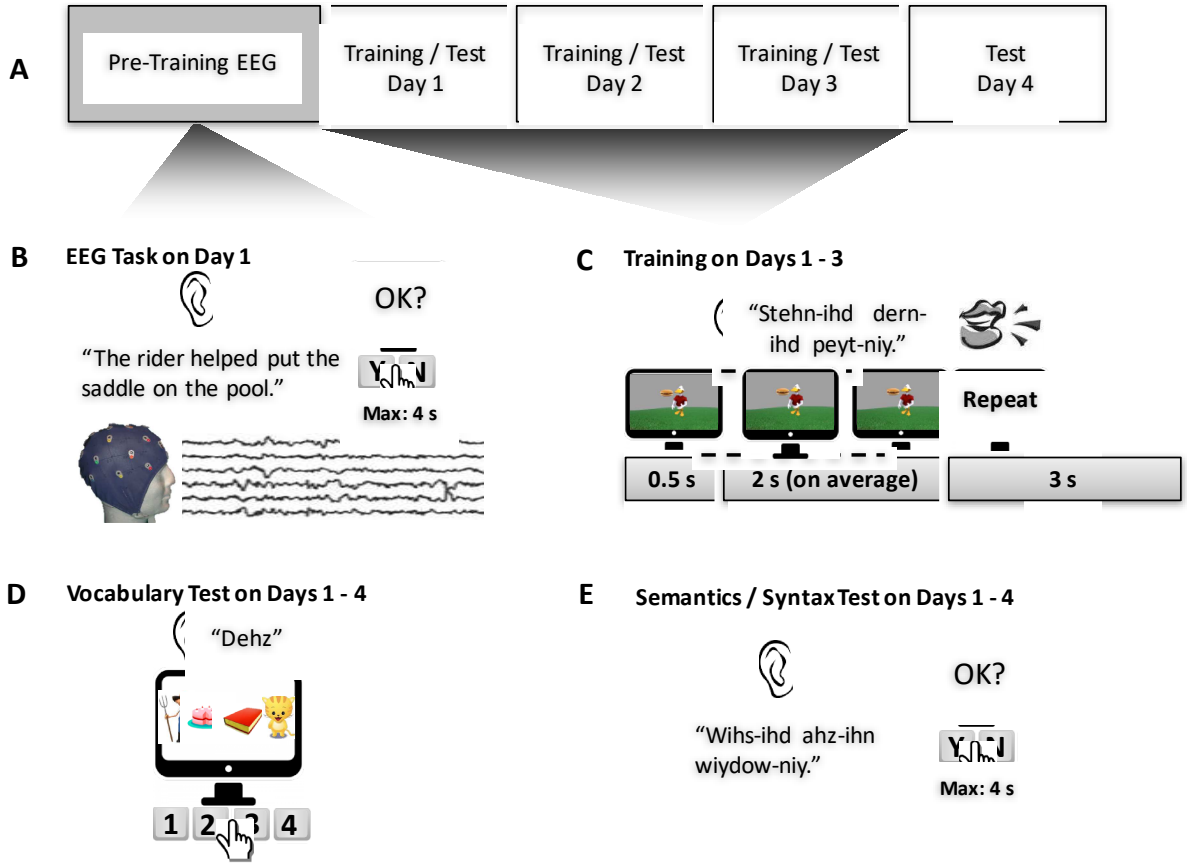


Fig. 1. Tasks and experimental design. **(A)** Participants completed a pre-training EEG session on Day 1 of the study, prior to training and testing on the miniature artificial language (MAL). Days 2 and 3 included MAL training and testing, and Day 4 included only MAL testing. The entire protocol spanned no more than six calendar days. **(B)** EEG was recorded while participants listened to English sentences and performed an acceptability judgment task. The unacceptable sentences contained either a semantic or a morphosyntactic violation. **(C)** On each MAL training trial, participants watched an animated clip, heard the pre-recorded sentence, and repeated it aloud. Each training sentence was presented 12 times over the course of three days. **(D)** Training was followed by a daily Vocabulary test. Participants heard a MAL word and had to select the corresponding picture. **(E)** Participants subsequently performed an auditory sentence acceptability judgment task in the MAL. The unacceptable sentences contained either a semantic or a morphosyntactic violation.

2.2 Participants

Thirty-eight native speakers of English (25 females and 13 males; mean age = 23.0 years, SD = 3.2 years, range = 19 – 32 years) participated. All participants had completed high school, were right-handed, and had no history of hearing impairment or neurological disorder. All

participants had average or above-average intelligence (age-based standard score ≥ 85), as measured by the Verbal (Verbal Knowledge, Riddles; mean = 116.2, SD = 13.8) and Nonverbal (Matrices; mean = 115.9, SD = 11.3) subtests of the Kaufman Brief Intelligence Test (KBIT-2; Kaufman & Kaufman, 2004). Participants had limited classroom-based foreign-language learning experience (mean = 2.43 years, SD = 1.84 years). All procedures were approved by the Massachusetts Institute of Technology Committee on the Use of Humans as Experimental Subjects, and informed written consent was obtained from all participants, who were compensated $30/hour for the EEG session and $20/hour for the behavioral training and testing sessions.

2.3 Day 1: EEG of native-language processing

2.3.1 Stimuli. One hundred twenty pairs of English sentences were taken from Tse et al. (2007), half of which were acceptable and half of which were unacceptable. Of the unacceptable sentences, half contained a semantic incongruency in the final word position, as in (a), and the other half contained a grammatical violation in the middle of the sentence, as in (b); critical words are underlined and correct forms are provided in parentheses.

(a) The rider helped put the saddle on the pool (horse).

(b) The silver plane took we (us) to paradise and back.

Each participant heard either the correct or the incorrect version of each sentence pair. In order to slightly decrease participants' expectations of an unacceptable sentence, the 120 target sentences were intermixed randomly with 30 filler sentences, of which 20 were correct and 10 contained semantically anomalous words in word positions 3-7. Each participant heard all 150 sentences evenly distributed over two blocks and arranged in a pseudorandom order such that no more than four consecutive trials had the same correct response.

Digital recordings of the sentences were made by a trained female native speaker of English in a double-walled acoustic chamber using a Shure WH20XLR microphone via a Roland UA-25EX sound card to a single channel with a sampling rate of 44.1 kHz. The intensity of each auditory recording was scaled to 73.0 dB. In order to ensure accurate event-related potential (ERP) time-locking to the critical word, one experienced researcher visualized each sentence waveform in Praat (Boersma & Weenink, 2011) while listening to the recording and marked the onset of the critical word. The onsets of the critical words were then second-coded by another researcher. The average discrepancy of onset time was 31 ms and there was no significant difference between the two coders (p = 0.79). We used the onsets marked by the first coder for ERP time-locking. On average, the critical word in the semantic conditions was uttered at 1604 ms (SD = 204 ms) after sentence onset. The critical word in the syntactic conditions was uttered at 1019 ms (SD = 225 ms). There was no significant difference between Acceptable and Unacceptable trials with regards to critical word onset time (two-sample $t(118) = 1.23$, $p = 0.22$, Cohen's $d = 0.22$) or syllable length (two-sample $t(118) = -0.19$, $p = 0.85$, Cohen's $d = 0.03$). The Acceptable and Unacceptable trials in the semantic conditions were also matched in the word frequency of the critical word, due to its influence on N400 amplitude (Van Petten & Kutas, 1990). Word frequency did not differ for Acceptable (mean = 53.65, SD = 46.96 per million words) and Unacceptable trials (mean = 60.14, SD = 61.89 per million words (retrieved from the American English subtitle database $SUBTLEX_{WF}$, two-sample $t(58) = -0.46$, $p = 0.65$, Cohen's $d = -0.12$).

2.3.2 Procedure. On Day 1, participants sat in a dimly lit sound-attenuating booth in front of an LCD monitor. Participants were asked to decide whether a sentence was acceptable. They listened to the English sentences played binaurally over Etimotic Research ER2 headphones for a

total of 20 minutes. A visual prompt ("Is this sentence okay?") appeared after each sentence and participants were made to respond "Yes" or "No" with a button press within four seconds (Fig. 1B). After responding, participants were given visual feedback of "Good!", "No!", or "Too slow!" After the first half of the task, participants saw their average accuracy and were given the option to take a break. Stimulus presentation and behavioral response recording were performed using Psychtoolbox-3 (Brainard, 1997) for MATLAB.

2.3.3 EEG recording. The electroencephalogram (EEG) was recorded continuously during the sentence judgment task from 32 scalp sites, as well as two reference sites on the left and right mastoids, using the Biosemi ActiveTwo System (Biosemi B.V., Amsterdam). To detect blinks and lateral eye movements for later correction, electrooculogram (EOG) electrodes were placed at the infraorbital ridge of the left eye and the lateral canthus of the right eye. The EEG was recorded with a low-pass hardware filter with a half-power cutoff at 104 Hz. Offset values for each electrode were kept below 40 mV. The data were digitized at 512 Hz with 24 bits of resolution and saved to a computer along with event timing information for the onset of the critical word in each sentence.

2.4 Days 1-3: Artificial language training

2.4.1 Stimuli. The miniature artificial language (MAL) was designed to mimic real languages of the world. It was modeled after an artificial language described by Finn and colleagues (2013), and was similar to other artificial languages (Hudson Kam & Chang, 2009; Hudson Kam & Newport, 2005, 2009). The MAL was meaningful and designed to be productive, so that learners could be assessed with novel sentences on which they had not been trained. The language consisted of four transitive verbs, 30 concrete nouns, two noun-specific particles, and two verb-agreement suffixes. The MAL used American English phonemes, but had a grammar distinct

from that of English. As shown in example (c) and its translation equivalent in (d) below, the MAL used subject-object-verb word order. Noun membership was indicated by one of two noun suffixes (*-ihd* or *-ihn*), which is similar to the marking of gender in Romance languages, and there was subject-verb agreement such that the verb suffix *(-niy* or *-ahn*) depended on the subject noun suffix (*-ihd* or *-ihn*).

(c) Snehl-ihd noy-ihn peyt-niy

(d) Woman(-ihd) rice(-ihn) eat(-niy)

Digital audio recordings of 60 MAL training sentences were made as described above. Visual stimuli were also created for the MAL training. The 30 noun images were selected to be highly identifiable and edited with Adobe Photoshop CS6. Using Adobe Flash CS6, these cartoons were animated into scenes of 2.5-s duration. The two nouns appeared simultaneously, and then the verb was acted out (for the MAL verb "eat") or emphasized with arrows (for "be on top of" and "be under"). Auditory and visual information were merged in Adobe Pro CS6.

2.4.2 Procedure. Training sessions were designed to be naturalistic. The 60 training sentences were repeated six times each on Day 1 and three times each on Days 2 and 3. Sentence presentation order was pseudo-randomized. Because verbs pose a special challenge in language learning (Golinkoff & Hirsh-Pasek, 2008), "verb sets," in which 3 or 4 consecutive sentences used the same verb in combination with different nouns, were preserved. Each participant saw the same sentences in the same order on each day, and had the opportunity to take a break every 15 minutes. For each training trial, participants watched an animated video scene, listened to a pre-recorded MAL sentence describing the scene, and repeated the sentence aloud (Fig. 1C). No feedback was provided during training. Training lasted approximately 30 to 45 minutes each day.

2.5 Days 1-4: Artificial language proficiency tests

Two proficiency tests were administered: a vocabulary forced choice task followed by a sentence acceptability task. Proficiency scores from Day 1 and Day 4 served as measures of initial learning and attainment, respectively.

2.5.1 Vocabulary. Vocabulary learning was tested by means of a four-alternative forced choice procedure in which the participants heard a MAL word and had to select one of four pictures that best matched the word (Fig. 1D). The pictures in the test were the same ones that had been used for training. The three picture foils for each test item were chosen randomly from the remaining 29 nouns. All 30 nouns were tested each day and presented in a random order each time, with the three picture foils consistent across subjects and sessions. No feedback was given.

2.5.2 Sentence acceptability judgment. Knowledge of MAL semantics and syntax was tested by means of a sentence acceptability judgment task (Fig. 1E). Participants heard novel, untrained MAL sentences and indicated whether each was acceptable or not in the same procedure as in English EEG task. However, in order to minimize learning during the test, feedback was not given.

Test sentences consisted of 160 novel sentence pairs, each containing one unacceptable sentence and its matched acceptable counterpart. Two lists, each containing 80 acceptable and 80 unacceptable sentences, were created by splitting up each pair, so that each participant heard only either the acceptable or the unacceptable form of each pair on each day. List 1 was tested on Days 1 and 3, and List 2 was tested on Days 2 and 4. Within each list, half of the items were flagged for semantic analysis and half for syntactic analysis, arranged in a pseudorandom order such that no more than four consecutive trials had the same correct response.

Semantics. Unacceptable sentences flagged for semantics contained a semantic incongruency in final word position. As demonstrated by English translation equivalents of MAL test items, either the verb was incongruent (e: A phone cannot eat a child), or the word order created an impossible situation (f: An apple cannot have a farmer). The critical word is underlined and the matched acceptable version is supplied in parentheses.

(e) Phone child <u>eat</u> (Phone child is under)

(f) Apple farmer <u>has</u> (Farmer apple has)

As knowledge of MAL lexical items is necessary for identifying these semantic violations, we hypothesized that semantic judgment accuracy would reflect individual differences in MAL Vocabulary knowledge and would be predicted by sensitivity to semantics (the N400) in the native language.

Syntax. Unacceptable sentences flagged for syntax contained a morphosyntactic violation, either a noun class mismatch (g) or a verb agreement violation (h). In (g), the MAL noun *rice* takes the suffix *-ihn*, not *-ihd*. In (h), the MAL verb *eat* must take the suffix *-niy* in order to agree with the subject *baby-ihd*.

(g) Student rice-<u>ihd</u> (rice-ihn) eat

(h) Baby-ihd cake eat-<u>ahn</u> (eat-niy)

Strictly speaking, only knowledge of structural co-occurrence information, and not of word meanings, is necessary for judging these sentences. Therefore, we hypothesized that syntactic judgment accuracy would be predicted by sensitivity to syntax (the P600) in the native language, rather than by the N400.

Thus, on each day, participants were tested on 160 MAL sentences, including 40 that were semantically unacceptable and 40 that were syntactically unacceptable. Participants were given the option to take a break after the first half of the task.

2.6 EEG Data Analysis

2.6.1 EEG Preprocessing. Electrophysiological data were analyzed using the EEGLAB 13.4.4b toolbox (Delorme & Makeig, 2004) and ERPLAB 4.0.3.1 (Lopez-Calderon & Luck, 2014). Scalp electrodes were referenced offline to the average of the two mastoids. The EEG was filtered at 0.1 Hz and 40 Hz and baselined on the 100 ms before critical word onset. We observed slow drift over the course of data acquisition in some participants. Although using a greater high-pass filter may help attenuate low-frequency noise, it has been shown that filter settings greater than 0.1 Hz can systematically introduce temporal distortions and spurious peaks in ERP waveforms (Acunzo, MacKenzie, & van Rossum, 2012; Tanner & Luck, 2016). We chose to preprocess the data without further filtering. Ocular artifacts were identified and removed using independent component analysis (Jung et al., 2000). Trials were removed from analysis if the peak-to-peak voltage between 100 ms pre-stimulus and 1500 ms post-stimulus exceeded 150 μV for any of the 32 EEG channels. On average, only 4.65% of trials (5.6 out of 120 critical trials) were removed from each participant's data. Acceptable and Unacceptable conditions did not differ significantly in the number of trials remaining after artifact rejection ($t(37) = 0.95$, $p = 0.35$). For each trial, ERPs were computed at each electrode, time-locked to the onset of the critical word in the utterance or its correct counterpart. Due to the high accuracy on the task (Section 3.1) there were too few incorrect trials for analysis, so all trials free of artifact were included in the ERP averages to maximize statistical power.

2.6.2 ERP Analysis. ERP components related to the processing of semantic and syntactic information in English were investigated separately. Selection of time windows of interest was guided by the substantial literature on the N400 and P600, so that our results could be interpreted within existing frameworks. We also verified the continuity of statistically significant effects within the *a priori* time windows with mass univariate analyses (Groppe et al., 2011). The window between 250 and 500 ms was selected *a priori* for the semantic congruency effect (Fig. 2). Reliable differences between ERPs elicited by semantically acceptable and unacceptable English sentences were determined by a repeated-measures, two-tailed permutation test based on the tmax statistic (Blair & Karniski, 1993) using a family-wise alpha level of 0.05. In order to identify all significant effects of semantic violation, all 32 scalp electrodes and all 513 time points between 0 and 1000 ms were included in the test. The distribution of the null hypothesis was estimated by 2500 random permutations of the data. Results indicated a widespread negative deflection peaking at approximately 400 ms. The earliest and latest time points with at least two significant channels were 244 and 547 ms. The results of the tmax test with the same parameters performed on a single, representative channel (Cz) are indicated by the blue bar on its waveform in Figure 2. These results are consistent with the literature on the N400 (Kutas & Hillyard, 1984; Kutas & Federmeier, 2011).

Two time windows (250-500 ms and 500-1000 ms) were selected *a priori* to examine the grammaticality effect (Fig. 3). While the 500-1000-ms window serves as the classic window of interest for P600 (Hagoort et al., 1993; Osterhout & Holcomb, 1992), examining the scalp distribution and voltage polarity of the effect between 250 and 500 ms would address the existence of the anterior negativity, the early P600, or the N400. An identical univariate test procedure was followed for syntactic processing. The tmax test revealed, similarly, a broad

positive deflection with significant effects in at least two channels beginning as early as 170 ms. Effects were mostly centro-parietal and lasted until 1000 ms. Representative tmax results from channel Pz are indicated with a blue bar in Figure 3. These results confirmed an overall positivity in both time windows.

ERP mean amplitudes were submitted to a repeated-measures ANOVA with factors of anteriority (anterior/posterior), laterality (left/central/right), and condition (acceptable/unacceptable), conducted separately for semantic and syntactic stimuli. Greenhouse–Geisser correction (Greenhouse & Geisser, 1959) was applied to avoid Type I errors in the case of a violation of the sphericity assumption. The ERP measures were the difference amplitudes averaged across the channels with the largest effect size, determined by a significant interaction between channel location and acceptability. The N400 effect was defined as (Acceptable – Unacceptable) and the P600 effect was defined as (Unacceptable – Acceptable), such that the canonical electrophysiological response to each would yield a positive number.

2.6.3 Response Dominance. Although the canonical responses to semantic and syntactic violations are negative-going and positive-going deflections, respectively, individuals have been shown to vary in the magnitude and polarity of their electrophysiological responses (Osterhout, 1997; Tanner & Van Hell, 2014). Therefore, a response dominance index (RDI) was calculated for each individual to capture the robustness of the classic pattern across two typical time windows (250-500 ms ("early") for the N400 and 500-1000 ms ("late") for the P600) but within the same subset of electrodes (see Section 3.2, Fig. 2 and 3).

$$\text{Semantic RDI} = (\text{N400 effect} - \text{semantic positivity effect})/\sqrt{2} =$$

$$((\text{Early}_{\text{Acceptable}} - \text{Early}_{\text{Unacceptable}}) - (\text{Late}_{\text{Unacceptable}} - \text{Late}_{\text{Acceptable}}))/\sqrt{2}$$

A positive value for the Semantic RDI indicates a typical N400 response, while a negative value indicates a midline late positivity to a semantic violation.

$$\text{Syntactic RDI} = (\text{P600 effect} - \text{syntactic negativity effect})/\sqrt{2} =$$

$$((\text{Late}_{\text{Unacceptable}} - \text{Late}_{\text{Acceptable}}) - (\text{Early}_{\text{Acceptable}} - \text{Early}_{\text{Unacceptable}}))/\sqrt{2}$$

A positive value for the Syntactic RDI indicates a typical P600 response, while a negative value indicates a posterior early negativity to a syntactic violation.

2.6.4 Brain-behavior correlations. As all proficiency measures were bounded between 0 and 1, Spearman correlations were conducted between three ERP measures (N400, early P600, and late P600) and five language proficiency measures, controlling *a priori* for individuals' age (Birdsong & Molis, 2001; Hakuta et al., 2003), gender (Nyikos, 1990), and nonverbal IQ (Brooks et al., 2006; Kaufman et al., 2009). Due to the strong correlation between Day 1 Vocabulary learning and verbal IQ (Spearman's $\rho = 0.60$, $p = 2.6\text{x}10^{-4}$), we elected not to eliminate this important source of variability. The five language proficiency measures include three initial learning measures (percent correct on Vocabulary, Semantics, and Syntax on Day 1) and two attainment measures (percent correct on Semantics and Syntax on Day 4). We did not include Vocabulary accuracy on Day 4 due to the fact that learners' performance had reached ceiling (Section 3.3). Because the N400 effect and P600 effect formulae were designed to yield positive numbers, we hypothesized that brain-behavior correlations would be positive, with greater N400 and P600 magnitudes serving as two aptitude markers for successful novel-language learning. Therefore, we performed one-tailed correlations, false-detection rate corrected for multiple comparisons (FDR < 0.05).

2.6.5 Ten-fold cross-validation. In order to test whether the ERP components help predict language-learning ability at the individual level, we validated significant correlations with 10-

fold cross-validation analysis. First we performed linear regression analysis on 90% of the participants as a training sample (N = 34), leaving out 10% of the participants (N = 4) to be tested. Training and testing participants were resampled 10 times. As a result, each participant was given a predicted score based on an independent training sample. The performance of each prediction model is reported in two measures: 1) the correlation coefficient between the predicted score and the actual score, and 2) the root mean square error (RMSE), measuring the normalized deviance of prediction error. The analysis was conducted with the DAAG package (Maindonald & Braun, 2014) in R 3.2.4 (R Core Team, 2014).

3. RESULTS

3.1 Native-language sentence acceptability task

Participants were accurate, and significantly better than chance, at judging the semantic (mean = 0.95, SD = 0.03, Wilcoxon signed rank test $V = 741$, $p = 6.60\text{x}10^{-8}$) and syntactic (mean = 0.97, SD = 0.03, Wilcoxon signed rank test $V = 741$, $p = 6.59\text{x}10^{-8}$) acceptability of sentences in English.

3.2 Native-language event-related potentials

We observed an overall negative shift in response to semantic incongruency in the 250-500-ms time window ($F(1, 37) = 56.01$, $p = 8\text{x}10^{-9}$). There was a significant interaction of laterality and semantic acceptability ($F(2, 74) = 14.93$, $p = 4\text{x}10^{-6}$), and Mauchly's test indicated that the assumption of sphericity had not been violated ($W = 0.99$, $p = 0.913$). There was no significant interaction of acceptability with anteriority ($F(1, 37) = 0.09$, $p = 0.77$). The magnitude of the semantic acceptability effect was greater in the center electrodes than it was in the

peripheral electrodes (Table S2). We therefore focused on eight center channels stretching from anterior to posterior (Fz, FC1, FC2, Cz, CP1, CP2, Pz, Oz) as the locus of the N400 effect (Fig. 2 & S1).

For syntactic processing, in the 250-500 ms time window, there was an overall positive shift in response to grammatical violations ($F(1, 37) = 7.03$, $p = 0.012$). There were two significant interactions with this effect: anteriority x acceptability ($F(1, 37) = 8.66$, $p = 0.006$) and laterality x acceptability ($F(2, 74) = 6.96$, $p = 0.002$; Mauchly's test was non-significant, $W = 0.99$, $p = 0.81$). The magnitude of the syntactic acceptability effect was greater in center as compared to peripheral electrodes (Table S2), and was greater in posterior as compared to anterior electrodes (Table S3). We therefore chose the posterior center electrodes (Cz, CP1, CP2, Pz, and Oz) as the locus of the early P600 effect (Fig. 3 & S2).

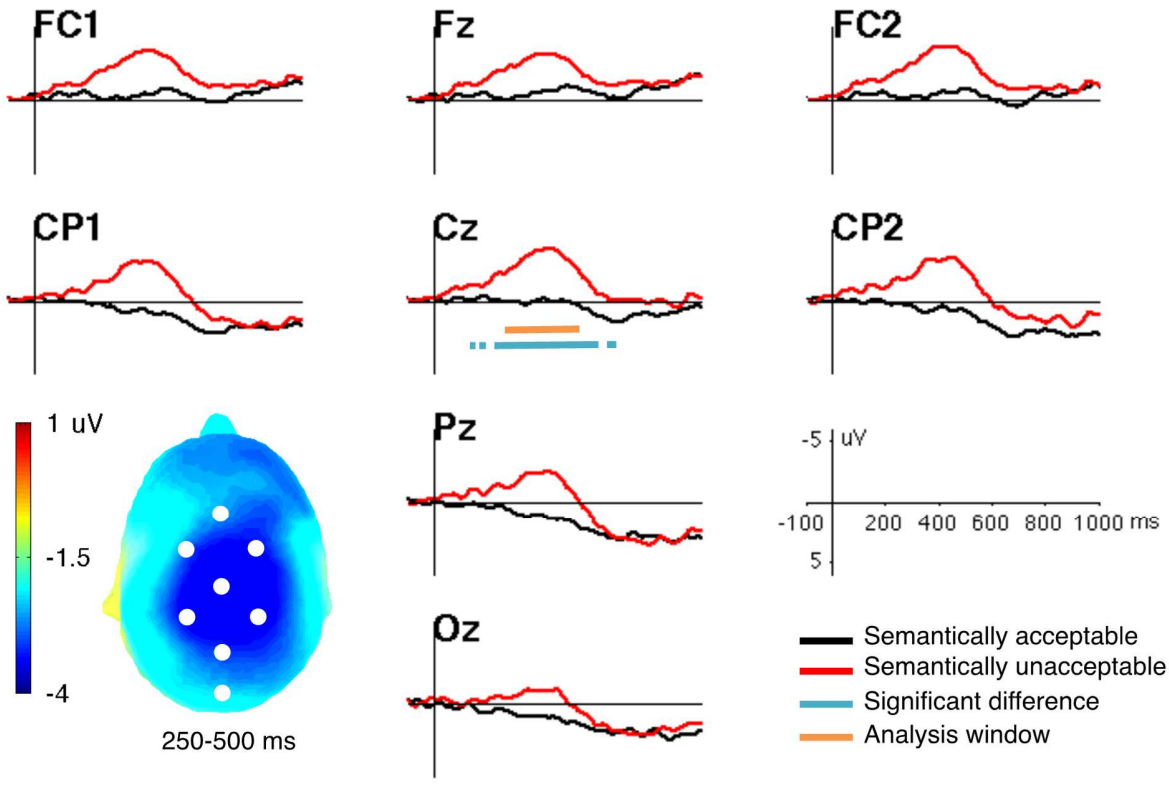


Fig. 2. Waveforms recorded in response to semantically acceptable and unacceptable sentences in English, with channel locations on the scalp indicated by white circles. Scalp plot shows the mean amplitude of the N400 effect over the 250-500-ms analysis window. Analysis window (250-500 ms) and all time points at which acceptable and unacceptable sentences differ significantly are indicated by bars on a representative channel, Cz.

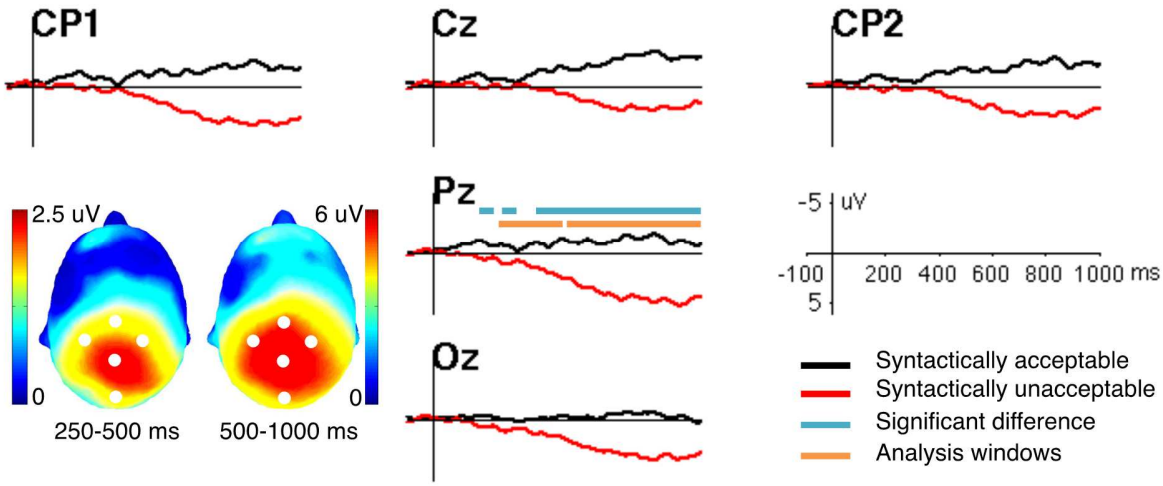


Fig. 3. Waveforms recorded in response to syntactically acceptable and unacceptable sentences in English, with channel locations on the scalp indicated by white circles. Scalp plot shows the mean amplitude of the P600 effect over the 250-1000-ms analysis window. Two analysis windows (early P600: 250-500 ms; late P600: 500-1000 ms) and all time points at which acceptable and unacceptable sentences differ significantly are indicated by bars on a representative channel, Pz.

Similarly, in the 500-1000-ms time window, there was a main effect of syntactic acceptability ($F(1, 37) = 29.89$, $p = 3\text{x}10^{-6}$), as well as two significant interactions, anteriority x acceptability ($F(1, 37) = 21.88$, $p = 4\text{x}10^{-5}$) and laterality x acceptability ($F(2, 74) = 15.98$, $p = 2\text{x}10^{-6}$; Mauchly's test was non-significant, $W = 0.96$, $p = 0.48$). The same five posterior center channels were chosen for this later P600 time window (Fig. 3 & S2, Table S2 & S3). There was no evidence of a left anterior negativity (LAN, 250-500 ms, $F(1, 37) = 0.378$, $p = 0.542$, Table S4) or an anterior negativity (AN, 250-500 ms, $F(1,37) = 1.241$, $p = 0.272$, Table S3). However,

a subset of participants (n = 9) showed syntactic negativity (Fig. 7B) at both anterior and posterior sites.[1]

3.3 Miniature artificial language (MAL) learning

Participants acquired a substantial amount of the MAL vocabulary after only one day of training (mean = 0.82, range = 0.33 – 1.00). The majority (76%) of participants achieved 100% accuracy by Day 4 (mean = 0.99, range = 0.87 – 1.00). A permutation test estimated by the Monte Carlo method showed significant Vocabulary growth from Day 1 to Day 4 (Fig. 4, $p = 5\text{x}10^{-4}$).

Compared with Vocabulary, participants' analysis of semantic and syntactic information in MAL sentences revealed slower growth and greater individual variability. Because accuracy on the two subtypes of semantic violations (verb, word order) was correlated on both Day 1 and Day 4, as was accuracy on the two types of syntactic violations (noun class, verb agreement) (p's ≤ 0.037), we elected to use overall accuracy on semantic judgments ("Semantics") and syntactic judgments ("Syntax") as outcome measures. On Day 1, participants were slightly better than chance when making judgments about semantic congruency (mean = 0.58, range = 0.29 – 0.94, Wilcoxon signed rank test $V = 487.5$, $p = 0.001$), but not when making judgments about grammaticality (mean = 0.32, range = 0.39 – 0.74, Wilcoxon signed rank test $V = 341.5$, $p = 0.46$). On Day 4, participants were able to judge both Semantics (mean = 0.64, range = 0.42 – 0.98, Wilcoxon signed rank test $V = 560$, $p = 6.07\text{x}10^{-5}$) and Syntax (mean = 0.66, range = 0.38 – 0.98, Wilcoxon signed rank test $V = 667.5$, $p = 1.67\text{x}10^{-5}$) reliably above the 50% chance level.

[1] We tested the existence of the LAN within this subset of participants. The negativity in the five left anterior channels (Fig. S2) was statistically reliable ($F(1,8) = 5.37$, $p = 0.049$), and was numerically larger than that in the posterior channels, albeit with an insignificant interaction between acceptability and anteriority ($F(1,8) = 1.15$; $p = 0.32$).

Permutation tests revealed significant improvements in Semantics (p = 0.04998) and Syntax (p = $5x10^{-4}$) from Day 1 to Day 4 (Fig. 4).

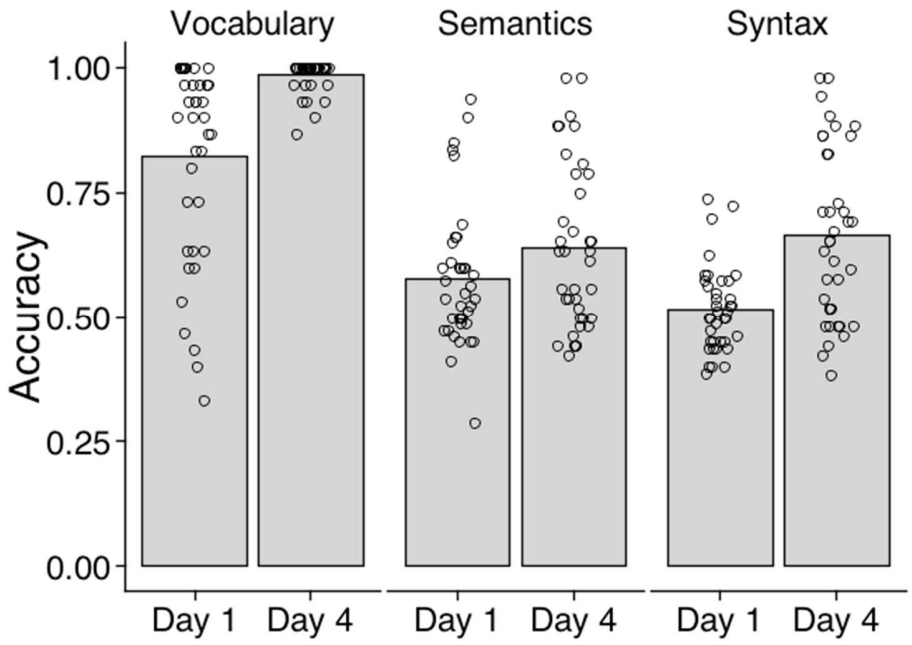


Fig. 4. Initial learning and attainment of the MAL. Vocabulary was measured by a forced-choice word-to-picture matching test (chance performance = 0.25). Semantic and syntactic knowledge were measured by sentence acceptability judgment (chance = 0.5). Each circle represents one participant.

3.4 Relationships between native-language ERPs and artificial-language learning

3.4.1 Relationship between semantic and syntactic learning

Vocabulary learning on Day 1 was positively associated with both Semantics and Syntax on both Day 1 and Day 4. We also observed significant correlations within learning domains. Initial learning of Semantics and Syntax was correlated with attainment of Semantics and Syntax, respectively (Table 1).

3.4.2 Neural sensitivity to semantic and syntactic information in the native language

The magnitude of the N400 effect was not correlated with the magnitude of the early P600 nor with the magnitude of the late P600 effect (Table 1).

3.4.3 ERP markers of native language processing predict MAL learning success

A larger English N400 effect was associated with better initial learning of Vocabulary, Semantics, and Syntax and marginally associated with better attainment of Semantics and Syntax (Fig. 6A, Table 1). In order to test whether the N400 predicted initial learning of Semantics and Syntax through successful acquisition of Vocabulary, we conducted structural equation modeling by using the *lavaan* R package (Rosseel, 2012). The analyses revealed non-significant direct associations between the N400 and Semantic learning and between the N400 and Syntax learning on Day 1. The N400 effect size was, however, indirectly associated with Semantics learning (Model A: $\chi^2 = 15.18$, $p = 0.002$) and Syntax learning (Model B: $\chi^2 = 9.43$, $p = 0.024$) through the mediation of Day 1 Vocabulary (Fig. 5). Based on the Akaike information criterion (AIC; Akaike, 1974), both models were better fits than the corresponding models without Vocabulary as the latent variable.

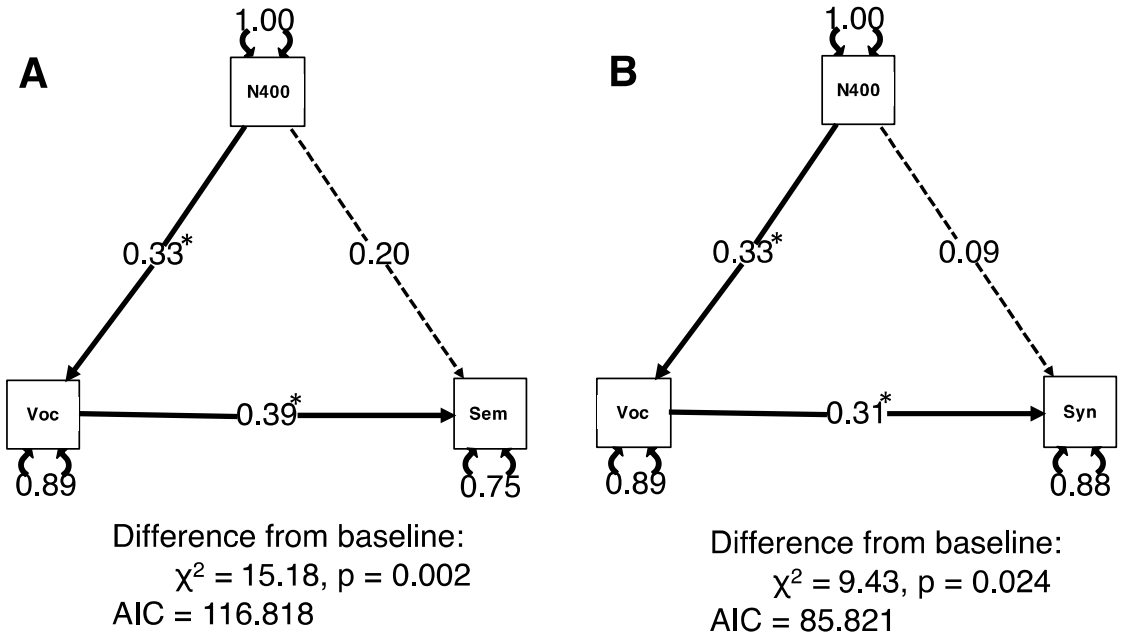


Fig. 5. Structural models of the relations between native-language N400 magnitude and language learning aptitudes. Voc: Vocabulary Day 1; Sem: Semantics Day 1; Syn: Syntax Day 1; *: significant path coefficients ($p < 0.050$).

A larger early P600 response was significantly associated with better initial Syntax learning (Fig. 6B). A larger late P600 was marginally associated with better initial Syntax learning (Table 1). There was no reliable correlation between the LAN or the AN and any learning outcome (Table S5).

Ten-fold cross-validation confirmed the relationships between English ERP predictors and MAL learning outcomes at the individual level. The N400 effect predicted Day 1 Vocabulary ($r = 0.34$, RMSE = 0.188). In combination, the N400 effect and participants' Day 1 Vocabulary predicted initial Semantics ($r = 0.38$, RMSE = 0.129). The early P600 alone predicted initial Syntax ($r = 0.33$, RMSE = 0.0804).

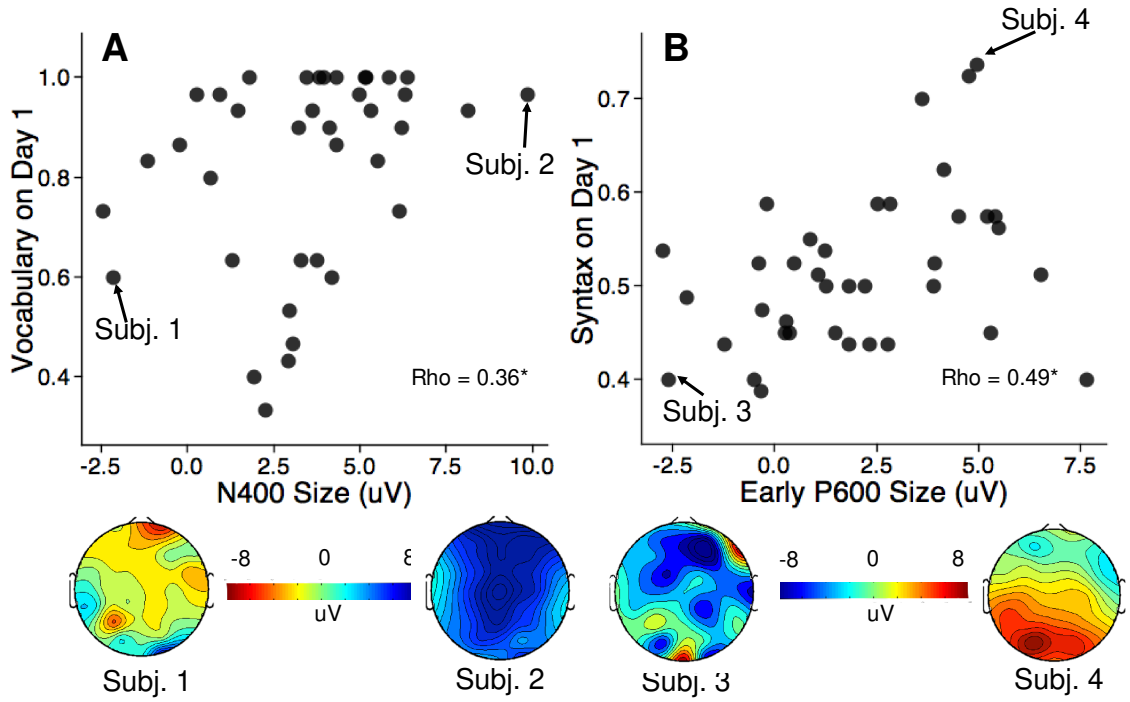


Fig. 6. Scatter plots between native-language ERPs and novel-language learning outcomes. **(A)** A larger N400 effect in response to semantic anomalies in English was associated with better Vocabulary learning on Day 1. **(B)** A larger early P600 effect in response to syntactic anomalies in English was associated with better Syntax learning on Day 1. Scalp distribution maps below the scatter plots depict four representative N400 and early P600 responses on both ends of the spectrum.

Table 1. Correlation matrix of 3 native-language-elicited ERP components and 5 MAL proficiency measures, controlling for gender, age, and nonverbal IQ. Estimates of Spearman's rho (ρ) are above the diagonal and uncorrected one-tailed p-values are below.

	N400	Early P600	Late P600	Vocab Day 1	Semantics Day 1	Semantics Day 4	Syntax Day 1	Syntax Day 4
N400		-0.081	-0.097	0.416*	0.358*	0.324†	0.355*	0.324†
Early P600	0.680		0.791*	-0.043	0.104	-0.071	0.489*	0.256
Late P600	0.712	0.000*		-0.114	-0.005	-0.171	0.275	0.202
Vocab Day 1	0.004*	0.598	0.745		0.521*	0.483*	0.316†	0.550*
Semantics Day 1	0.014*	0.273	0.512	0.000*		0.346*	0.222	0.170
Semantics Day 4	0.025†	0.658	0.840	0.001*	0.017*		0.159	0.096
Syntax Day 1	0.015*	0.001*	0.050	0.028†	0.096	0.178		0.388*
Syntax Day 4	0.024†	0.064	0.118	0.000*	0.160	0.290	0.008*	

* Significant after FDR correction (FDR < 0.050)
† Marginal ($p \leq 0.060$) after FDR correction
Relationships between semantic and syntactic learning
Relationships between N400 and P600 effects
Relationships between N400, P600, and MAL proficiency

3.4.4 Testing the double dissociation pattern in the relationships between native-language ERPs and language-learning outcomes

Based on pairwise correlation analysis, only the relationships between ERPs and the initial learning of Vocabulary, Semantics and Syntax remained significant after correcting for multiple comparisons (Table 1). Moreover, structural equation modeling revealed that the N400-Semantics relationship was mediated by the initial learning of Vocabulary (Fig. 5). Therefore, in follow-up analyses, we asked whether two specific learning outcomes, initial Vocabulary and initial Syntax, representing sound-to-meaning mapping and structural learning respectively, were better predicted by one ERP marker or the other. Because accuracy on these measures was correlated, we controlled for the contributions of each. Initial Vocabulary was significantly predicted by the N400, but not by the early P600. Initial Syntax was marginally predicted by the N400 and significantly predicted by the early P600 (Table 2). The N400-Vocabulary relationship was significantly stronger than the early P600-Vocabulary relationship, and the early P600-Syntax relationship was marginally stronger than the N400-Syntax relationship (Table 2).

We further examined whether each ERP marker was a better predictor of one learning outcome than the other learning outcome. The N400 significantly predicted both initial Vocabulary and initial Syntax, with neither correlation stronger than the other (Table 3). However, structural equation modeling confirmed that the N400 predicted initial Syntax via Vocabulary (Figure 5B). The early P600 was a significantly better predictor of initial Syntax than of initial Vocabulary (Table 3).

Table 2. Comparison of ERP predictors of MAL learning using estimates of Spearman's rho (ρ). "N400 vs. Early P600" refers to Steiger's (1980) test of the relative strength of two correlations. Reading within each column, the N400 is a statistically stronger predictor than the early P600 of Vocabulary, and the early P600 is a marginally stronger predictor than the N400 of Syntax.

	Vocab Day 1, Controlling for Syntax Day 1	**Syntax Day 1, Controlling for Vocab Day 1**
N400	ρ = 0.342, p = 0.020	ρ = 0.259, p = 0.065
Early P600	ρ = -0.239, p = 0.918	ρ = 0.530, p = 0.000
N400 vs. Early P600	z = 2.62, p = 0.004	z = -1.36, p = 0.087

Table 3. Comparison of MAL learning outcomes predicted by ERPs using estimates of Spearman's rho (ρ). "Vocab vs. Syntax" refers to Steiger's (1980) test of the relative strength of two correlations. Reading within each row, the N400 is not a significantly better predictor of Vocabulary or Syntax, but the early P600 is a significantly better predictor of Syntax than of Vocabulary.

	Vocab Day 1	**Syntax Day 1**	**Vocab vs. Syntax**
N400	ρ = 0.416, p = 0.004	ρ = 0.355, p = 0.015	z = 0.345, p = 0.365
Early P600	ρ = -0.043, p = 0.598	ρ = 0.489, p = 0.001	z = -2.875, p = 0.002

3.5 Relationship between English ERP response dominance and MAL learning

Participants exhibited large variability in their native-language ERPs, ranging from a typical N400 to a positive deflection in response to semantic anomalies, and from a typical P600 to a negative deflection in response to syntactic anomalies (Fig. 7). In order to further validate the specificity of the relationship between ERP components and MAL learning, we examined partial Spearman correlations between the Semantic and Syntactic response dominance indices (RDI) and MAL learning, while controlling for gender, age, and nonverbal intelligence (Table 4). Similar to the pairwise correlation results in Table 1, a more positive Semantic RDI (i.e., a more typical N400 response) was associated with better initial Vocabulary, Semantics attainment, and initial Syntax. A more positive Syntactic RDI (i.e., a more typical P600 response) was associated only with better initial Syntax.

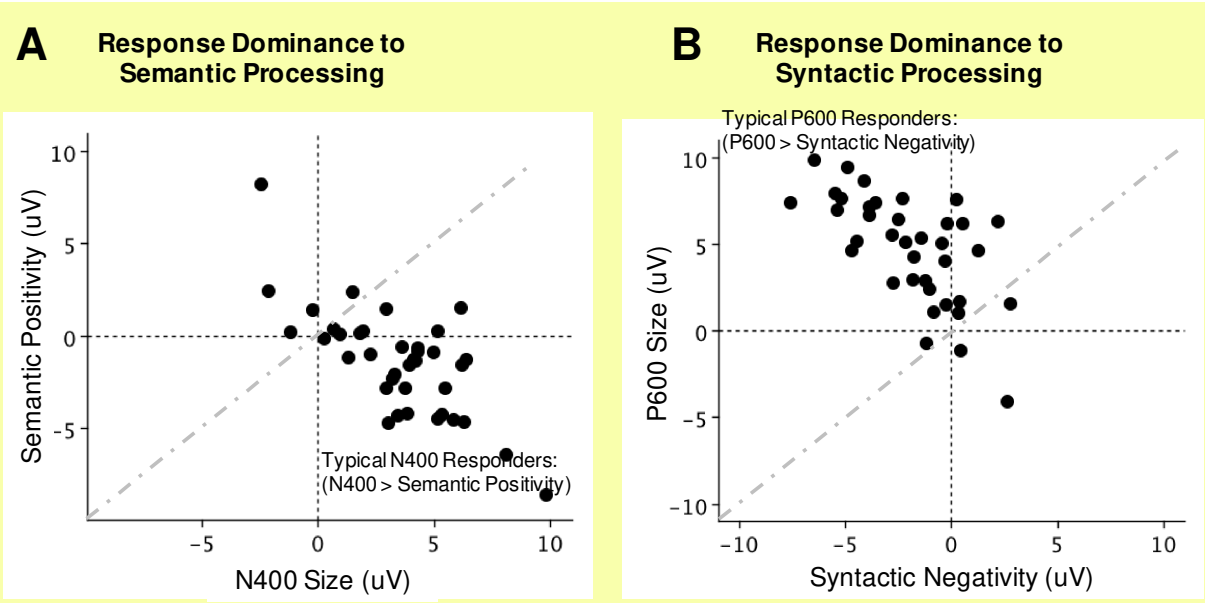


Fig. 7. Scatter plots reflecting individuals' response dominance across early (250-500 ms, x-axes) and late (500-1000 ms, y-axes) time windows for **(A)** semantic processing and **(B)** syntactic processing of the native language. In **(A)**, individuals in the bottom right quadrant show the canonical N400 to semantic anomalies. In **(B)**, individuals in the top left quadrant show the canonical P600 to syntactic violations.

Table 4. Correlations among response dominance indices (RDI) and MAL proficiency measures, controlling for gender, age, and nonverbal IQ. Estimates of Spearman's rho (ρ) and uncorrected one-tailed p-values are provided.

	Vocab Day 1	Semantics Day 1	Semantics Day 4	Syntax Day 1	Syntax Day 4
Semantic RDI	ρ = 0.351 p = 0.016	ρ = 0.267 p = 0.055	ρ = 0.334 p = 0.021	ρ = 0.296 p = 0.038	ρ = 0.197 p = 0.125
Syntactic RDI	ρ = -0.105 p = 0.728	ρ = 0.050 p = 0.386	ρ = -0.152 p = 0.811	ρ = 0.399 p = 0.006	ρ = 0.225 p = 0.092

4. DISCUSSION

4.1 Overview of results

Despite being highly proficient in their native language, individuals exhibit substantial variability in their electrophysiological responses during native-language processing (Osterhout, 1997; Tanner & Van Hell, 2014). Meanwhile, adults show a wide range of outcomes when learning a novel language. Here, for the first time, we showed that the extent to which typical N400 and P600 responses are elicited by anomalies in the native language predicted individuals' success in acquiring, respectively, novel lexical-semantic and syntactic information in an immersive learning paradigm. Moreover, we demonstrated a double-dissociation pattern in which novel word learning was more strongly predicted by the ERP marker of native-language semantic processing than that of syntactic processing, while novel morphosyntactic learning trended towards being better predicted by the ERP marker of native-language syntactic processing than that of semantic processing. We used a ten-fold cross-validation procedure to confirm that these native-language ERP measures predicted specific language-learning outcomes at the individual level.

4.2 N400 and P600 responses during auditory sentence comprehension in English

Across 38 native English speakers, a significant N400 effect of semantic congruency was observed in the 250-500-ms time window. A significant P600 effect of grammaticality was observed in both early (250-500 ms) and late (500-1000 ms) time windows.

The group average ERP responses to morphosyntactic violations in our sample revealed two notable differences as compared to previous literature. First, we did not observe the LAN reported in earlier ERP studies (Friederici, Hahn, & Mecklinger, 1996; Friederici & Frisch,

2000; Coulson et al., 1998; Münte, Heinze & Mangun, 1993; Münte et al., 1997; Gunter, Stowe, & Mulder, 1997). The LAN/AN is often interpreted as reflecting neural sensitivity to morphosyntactic errors, and together with a later positivity (P600), may contribute to a biphasic LAN-P600 pattern (Steinhauer & Drury, 2012). Nevertheless, elicitation of the LAN/AN has not been consistently reported across studies at the group level. Other studies investigating morphosyntactic violations reported an absence of the LAN/AN (Osterhout et al., 1996; Allen, Badecker & Osterhout, 2003; Hagoort & Brown, 2000). Evidence from recent individual-difference studies (Tanner & Van Hell, 2014) and the current findings revealed that the existence or absence of the LAN/AN could be driven by subgroups of participants. As we discuss below, participants displayed a continuum of responses ranging from a syntactic negativity to the canonical P600 in response to morphosyntactic violations. In our sample, a subset of participants showed signs of the biphasic LAN-P600 pattern that is consistent with many group analysis results reported previously.

Second, in the 250-500-ms time window, we observed a positive shift (the early P600) instead. The P600 effect is generally long lasting, and it is not uncommon for researchers to define two or even three time windows for its analysis. A review of 29 ERP studies of grammatical agreement in written language showed that effects are commonly reported in the early (roughly 500-700 ms) or late (roughly post-700 ms) time windows, and sometimes in both (Molinaro et al., 2011). However, modality may play a role, as the P600 elicited by auditory presentation appears to emerge earlier (Osterhout & Holcomb, 1993; Patel et al., 1998). In the present study, we used auditory presentation and observed an early and sustained effect of grammaticality. ERPs to acceptable and unacceptable items began to diverge around 200 ms after critical word onset. Our syntactically unacceptable sentences all made use of an

inappropriate pronoun form (e.g., *we/us*; *I/me*). As aural comprehension proceeds phoneme by phoneme, some pronoun violations can be registered relatively close in time to the critical word's onset, unlike the verb agreement errors used in other P600 paradigms, which involve the final morpheme (e.g., *walk/walks*). Therefore, the early syntactic effect could be attributed to a relatively short processing time.

Brain responses during auditory sentence comprehension varied profoundly across individuals, even though listeners were highly accurate in judging the acceptability of the English sentences. For each individual, we computed a semantic response dominance index (Semantic RDI) and a syntactic response dominance index (Syntactic RDI) by combining the effect of acceptability across two time windows in order to account for possible individual differences not only in ERP magnitude, but also in latency and/or polarity. We replicated the finding of a response-dominance continuum during morphosyntactic processing (Fig. 7B; cf. Osterhout, 1997; Tanner & Van Hell, 2014). A similar continuum has also been reported in language learners (Morgan-Short et al., 2012b; Tanner et al., 2013; McLaughlin et al., 2010; Nakano et al., 2010). The appearance of the syntactic negativity in early L2 learners has been interpreted as immature syntactic processing that relies on word forms prior to the emergence of a P600 (McLaughlin et al., 2010; Osterhout, 1997). The present ERP findings indicate a comparable range of neural responses to morphosyntactic information even in highly proficient native-language speakers.

We extended this work with evidence of a similar continuum during semantic processing (Fig. 7A). Consistent with previous research (Kos et al., 2012), five participants showed a late positive component ("semantic P600") over posterior scalp sites between 500 and 1000 ms in

conjunction with the N400, while an even smaller subset showed the positive deflection alone, without the N400.

Importantly, within-subject N400 effect size and P600 effect size (in either time window) did not correlate with each other (Table 1). Therefore, it is not the case that there are "good responders" who are generally more sensitive to anomalies in the native language. These results are congruent with theories of language comprehension that postulate distinct neural substrates for semantic and syntactic processing (for reviews, see Hagoort & Indefrey, 2014; Osterhout et al., 2012).

4.3 Individual differences in language-learning aptitude

We observed a wide range of outcomes for young adults learning a miniature artificial language (MAL) in terms of initial learning (Day 1) and attainment (Day 4). The variability in Vocabulary learning was most profound on Day 1, as three training sessions were largely sufficient for individuals to learn the whole MAL lexicon. In contrast, sentence comprehension skills as reflected by Day 1 and Day 4 performance on the MAL sentence acceptability judgment task demonstrated, on average, slower and more limited growth. Vocabulary acquisition had a positive association with both semantic and syntactic sentence acceptability judgment accuracy (Table 1; Fig. 5). While there were no significant correlations between Semantics and Syntax, there was a trend on Day 1 for better performance on one to track with better performance on the other. We then investigated in detail whether individual learners have different capacities for learning distinct aspects of language.

4.4 Native-language N400 effect predicts vocabulary learning

A larger native-language N400 effect during auditory comprehension of semantically incongruent and congruent sentences in the 250-500-ms time window was significantly

associated with better initial learning of MAL Vocabulary, Semantics, and Syntax, and also with better attainment of Semantics and Syntax (marginal after FDR correction) (Table 1). Ten-fold cross-validation confirmed that the model was not overfit to the data: the N400 effect size predicted MAL Vocabulary learning and generalized to untrained folds of the data. In post-hoc analyses, we found that the N400 had an indirect relationship to initial Semantics and Syntax via initial Vocabulary learning (Fig. 5). As a predictor, the N400 had a stronger relationship to the unique aspects of Vocabulary learning than did the early P600 (Table 2). An additional, more holistic measure of native-language semantic processing, the Semantic RDI, was associated with initial Vocabulary and Semantics attainment (Table 4). This confirmed that N400 typicality is associated with effective learning of novel lexical-semantic information. Together, these results demonstrate a specific association between the native-language N400 and Vocabulary learning, which enables sentence comprehension.

The sentence acceptability judgment task we employed is prevalent in the psycholinguistic literature (e.g., Batterink & Neville, 2013; McLaughlin et al., 2010; Osterhout, 1997). In order to find a MAL sentence such as *Snehl-ihd noy-ihn peyt-niy* acceptable, one must integrate multiple types of linguistic knowledge, including lexical semantics (*snehl* = *woman*, *noy* = *rice*, *peyt* = *eat*), word order (subject-object-verb), and world knowledge (a woman can eat rice). Given a straightforward, fixed word order and intact world knowledge, vocabulary is likely to be the main driver of inter-subject variability, and indeed Vocabulary correlates with Semantics (Table 1). Structural equation modeling demonstrated that native-language semantic processing is related to MAL Semantics via the effective learning of Vocabulary.

We measured the N400 response as a well-characterized index of semantic integration, sensitive to the congruency of a word in sentence context (as in our English stimuli) and to

meaning in general (as evidenced by semantic priming, word-pseudoword comparisons, etc. (e.g., Bentin et al., 1985)). The N400 reflects stimulus-induced activity in long-term semantic memory and indexes semantic integration (Kutas & Federmeier, 2011). A larger N400 amplitude implies more elaborated semantic networks. Previous individual difference studies showed that N400 amplitude (and in some cases, latency) is sensitive to such factors as semantic memory in the native language (Federmeier et al., 2002; Perfetti et al., 2005) and, in a novel language, the degree of language proficiency (Batterink & Neville, 2013; Newman et al., 2012). In children, the N400 has been found to correlate with vocabulary size (Khalifian et al., 2015); interestingly, high vocabulary and verbal fluency seem to serve as protective factors for a subset of older adults whose N400 response pattern, like that of younger adults, evinces predictive facilitation of sentence processing (Federmeier et al., 2002).

MEG and fMRI have contributed to our understanding of the source(s) of the N400 component. MEG source estimation implicates left posterior temporal cortex in the generation of the N400 (Service et al., 2007). Left posterior temporal cortex is part of the lexical-semantic network (Hickok & Poeppel, 2007; Lau et al., 2008), and lesions to this area lead to a specific impairment in single word comprehension (Dronkers et al., 2004). A meta-analysis of studies contrasting responses to spoken words and pseudowords has also identified posterior temporal areas as part of a large left-hemisphere cluster sensitive to lexicality (Davis & Gaskell, 2009). In a MEG study in which participants learned the names and definitions of novel words, learning effects were noted in left temporal cortex only (Hultén et al., 2009). Given converging evidence that the N400 is generated in cortical areas that represent lexical semantics and support word learning, it is reasonable that the N400 effect we measured would predict Vocabulary learning.

4.5 Native-language P600 effect predicts syntactic learning

A positive-going effect elicited by English syntactic violations was measured in two time windows, 250-500 and 500-1000 ms. A larger magnitude of the early P600 effect was significantly associated with better initial MAL Syntax learning (Table 1). Ten-fold cross-validation confirmed that the model was not overfit to the data: the P600 effect size predicted initial Syntax learning and generalized to untrained folds of the data. In post-hoc analyses, we found that the early P600 was also related to initial Syntax learning above and beyond the contributions of Vocabulary knowledge, was a marginally stronger predictor than the N400 for initial Syntax (Table 2), and was a significantly stronger predictor of initial Syntax learning than of Vocabulary learning (Table 3). As an additional, more holistic measure of native-language syntactic processing, we calculated the Syntactic RDI and found that it was associated with superior initial learning as well as attainment of Syntax (Table 4). These results demonstrate a specific association between the native-language P600 and Syntax learning, which appears to be a process independent of Vocabulary learning.

Like Semantics, Syntax learning was measured by means of sentence acceptability judgments. However, because participants were tested on morphosyntactic violations of noun class or verb agreement, only the knowledge of structural co-occurrence was required. Specifically, using the distributional information embedded across multiple trials, participants had to learn that MAL nouns always occur with their particular suffixes (*-ihn* or *-ihd*), and that there is a non-adjacent dependency between the suffix of the first word (the subject) and the suffix of the last word (the verb). Strictly speaking, neither word meanings nor world knowledge were required in order to detect syntactic anomalies. However, knowledge of word meanings could also facilitate the learning of these syntactic rules, according to the theory of

semantic bootstrapping (Pinker, 1984). For example, knowing the mapping between the meaning and the sound of an animate noun in the MAL helps to assign its agent role in the sentence and further supports the acquisition of subject-verb agreement. Therefore, it is not surprising that Vocabulary and the N400 effect were related to syntactic learning. More importantly, however, the early P600 predicted initial Syntax performance above and beyond the contributions of the age, gender, nonverbal intelligence, and Vocabulary covariates. These data suggest that the early P600, which indexes individuals' sensitivity to native-language syntactic rules, predicts their learning of novel morphosyntactic rules.

The relationship we found between the native-language Syntactic RDI and morphosyntactic learning aptitude further supports the double-dissociation pattern. In our study, a few individuals showed a "syntactic negativity" over posterior sites rather than a P600 response to native-language syntactic anomalies and were relatively poor in learning novel morphosyntactic rules in the MAL. Our findings suggest that the native-language posterior syntactic negativity might serve as a marker for a sub-optimal mechanism for learning the structural co-occurrence patterns of a novel language.

Debates over the nature of the P600 have revolved around its relationship to the P300 family of components (Coulson, King & Kutas, 1998; Osterhout & Hagoort, 1999; Osterhout et al., 1996; Sassenhagen, Schlesewsky & Bornkessel-Schlesewsky, 2014) and its specificity to syntax or even language. Although syntactic violations reliably elicit P600s, it may be the case that the neural generator(s) of the P600 are simply pattern and sequence processors (Osterhout, Kim & Kuperberg, 2012) and that linguistic syntax robustly engages these processes. We found significant positive effects of morphosyntactic violation in both early and late time windows. That the earlier, but not the more typical, P600 *time window* was predictive of learning might

suggest that those whose P600 had an earlier onset were better learners. Alternatively, the long-lasting positivity might reflect the summation of an earlier *component* with a similar topography and the P600, with the earlier component bearing the true relationship to syntactic learning ability. Individual differences were captured in the more variable immediate diagnosis of morphosyntactic violation (Mecklinger et al., 1995; Patel et al., 1998), rather than in the less variable canonical P600 representing the structural reanalysis process (Kaan et al., 2000; Hagoort & Brown, 2000; Friederici, Hahne, & Mecklinger, 1996).The present data cannot adjudicate amongst these hypotheses. Nevertheless, it is the neural response in the earlier window/component that differentiates learners. Learners with greater initial sensitivity to morphosyntactic violations were better learners.

4.6 Contribution of native-language linguistic sensitivities

Individuals' language-learning success may be attributed to dissociable linguistic aptitudes. In the present study, greater neural sensitivity to particular aspects of English carried over into the learning of the MAL, predicting individuals' success in learning novel Vocabulary, Semantics, and Syntax. We found a double-dissociation pattern in the relationship between two ERP predictors (the markers of native-language semantic processing and syntactic processing) and two outcomes (vocabulary learning and morphosyntactic learning), albeit with some nuance. The early P600, an index of sensitivity to structural regularity, was uniquely related to individuals' initial learning of novel morphosyntactic rules, as opposed to initial learning of vocabulary. The N400, an index of the attempt to access semantic information, was more generally related to successful learning. However, the N400 was a stronger predictor of vocabulary learning than was the early P600, and was indirectly associated with semantic and syntactic learning via its relationship with vocabulary. Together, these results suggest that

individual differences in the native language are related to corresponding abilities in a novel language.

4.7 Contribution of verbal intelligence and domain-general cognition

Alternatively, instead of dissociable linguistic processing abilities, the N400 and P600 might index broader verbal intelligence and cognitive abilities, which in turn might underlie language-learning success. In our sample, the magnitude of the N400 effect correlated positively with verbal IQ, which measures crystalized knowledge of word meanings and relationships in the native language. Due to the N400's association with semantic access, not surprisingly, verbal IQ was also positively associated with initial MAL vocabulary learning, but not with other learning outcomes.

Variability in the N400 and P600 has previously been linked to individual differences in cognitive ability, such as working memory (WM) span. Readers with low WM span showed reduced sensitivity, as measured by the N400, to facilitating sentential or lexical context (Salisbury, 2004; Van Petten et al., 1997). Furthermore, individuals with low WM spans did not show a P600 to the disambiguating word in a complex sentence, whereas high-span individuals did (Vos & Friederici, 2003). When parsing sentences containing long-distance syntactic dependencies, low-span individuals were about 200 ms delayed, based on the longer latency of their P600 responses (Hestvik et al., 2012). However, although the N400 and P600 have both been associated with WM, the English sentence stimuli in our study were simple enough that they neither taxed WM nor revealed its variability. Therefore, our results do not provide evidence either for or against the role of WM in language learning.

4.8 Implications for novel language learning

Our results contribute to an understanding of the variability in adults' novel-language learning trajectories. Whereas native-language acquisition proceeds unidirectionally through a set of developmental milestones commensurate with brain growth (Sakai, 2005), novel-language learning in adults may be quite idiosyncratic. For example, novel-language literacy may precede speaking and aural comprehension, which is never the case with native languages. Individuals vary in the degree to which they attain proficiency in foreign-language phonology, semantics, and syntax, and their skills may be strikingly asymmetric. Here, we demonstrate a dissociation between semantics aptitude and syntax aptitude. Learners generally performed better on one or the other, and our ERP results show that greater neural sensitivity to particular aspects of English carried over into the learning of the MAL, predicting individuals' success in learning novel Vocabulary, Semantics, and Syntax. There is a tradeoff, of course, in predicting the learning of an artificial language rather than a natural language. On one hand, artificial languages allow experimenters to precisely elaborate or eliminate certain linguistic features and to control the amount and manner of learners' exposure. On the other hand, they lack the richness, complexity, and embeddedness of natural languages. Nevertheless, a learned artificial language and a native language may, under some conditions, elicit similar patterns of brain activity (Friederici, Steinhauer & Pfeifer, 2002; Morgan-Short et al., 2012b). Moreover, it has been shown that artificial-language performance correlates positively with second-language learning success, even after controlling for IQ (Ettlinger et al., 2015).

The predictive power of the N400 and P600 effects was particularly notable for the initial learning that occurred during a single, 45-minute immersive training session. We propose that initial learning is closely related to individuals' aptitude, while their attainment after several

sessions may be influenced by additional factors such as motivation, self-monitoring, or even sleep. One limitation of the current study is that we were not able to address the contributions of such variables in a quantitative manner. Nevertheless, aptitude may determine an individual's early perception of success and willingness to pursue further language study. The rate of second-language learning has been predicted by resting-state EEG power in particular frequency bands (Prat et al., 2016). That study and ours suggest that electrophysiology may offer insight into learning ability before language exposure begins.

The present study provides a novel perspective on individual differences as they demonstrate neural variation in, as well as aptitude for, language processing. We found that individuals who show *robust* and *distinct* neural responses to native-language semantic and syntactic probes also show greater sensitivity to the lexical and structural features, respectively, of a novel language. Thus, the native-language N400 and P600 constitute markers of "neural preparedness" for learning novel semantics and syntax.

ACKNOWLEDGEMENTS

We thank Carla Hudson Kam and Marc Ettlinger for their help in designing and creating the miniature artificial language, and Susan Garnsey for generously sharing the English sentence stimuli. We thank Kana Okano, Arim Choi Perrachione, Caitlin Tan, Callie McRee, and Tyler Perrachione for their assistance with data acquisition, as well as Atsushi Takahashi, Sheeba Arnold Anteraper, and Steven Shannon at the Athinoula A. Martinos Imaging Center at the McGovern Institute for Brain Research at MIT for technical support. We would also like to thank two anonymous reviewers for their helpful comments on this manuscript. The research is based upon work supported by the Office of the Director of National Intelligence (ODNI), Intelligence

Advanced Research Projects Activity (IARPA), and the Air Force Research Laboratory via contract #FA8650-12-C-7265. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the ODNI, IARPA, AFRL, or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright annotation thereon. SDB was supported by NIH Training Grant T32DC000038.

REFERENCES

Acunzo, D. J., MacKenzie, G., & van Rossum, M. C. (2012). Systematic biases in early ERP and ERF components as a result of high-pass filtering. *Journal of neuroscience methods*, *209*(1), 212-218.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE transactions on automatic control*, 19(6), 716-723.

Alexander, M. P., Naeser, M. A., & Palumbo, C. (1990). Broca's area aphasias: Aphasia after lesions including the frontal operculum. *Neurology*, *40*(2), 353–353.

Allen, M., Badecker, W., & Osterhout, L. (2003). Morphological analysis in sentence processing: An ERP study. *Language and Cognitive Processes*,*18*(4), 405-430.

Batterink, L., & Neville, H. J. (2013). ERPs Recorded during Early Second Language Exposure Predict Syntactic Learning. *Journal of Cognitive Neuroscience*, *25*(6), 936–951.

Benson, D. F. (1985). Aphasia. In K. Heilman & E. Valenstein (Eds.), *Clinical neuropsychology*. New Yrok: Oxford University Press.

Bentin, S., McCarthy, G., & Wood, C. C. (1985). Event-related potentials, lexical decision and semantic priming. *Electroencephalography and Clinical Neurophysiology*, *60*(4), 343–355.

Birdsong, D., & Molis, M. (2001). On the evidence for maturational constraints in second-language acquisition. Journal of Memory and language,44(2), 235-249.

Blair, R. C., & Karniski, W. (1993). An alternative method for significance testing of waveform difference potentials. *Psychophysiology*, *30*(5), 518–524.

Boersma, P., & Weenink, D. (2011). Praat: Doing phonetics by computer (Version 5.1. 09)[Computer program].

Brainard, D. (1997). The psychophysics toolbox. *Spatial Vision*, 433–436.

Brooks, P. J., Kempe, V., & Sionov, A. (2006). The role of learner and input variables in learning inflectional morphology. Applied Psycholinguistics,27(02), 185-209.

Carroll, J. B., Sapon, S. M., Reed, D. J., & Stansfield, C. W. (2010). *Modern Language Aptitude Test: Manual.* Rockville, MD: Second Language Testing Foundation.

Coulson, S., King, J. W., & Kutas, M. (1998). Expect the unexpected: Event-related brain response to morphosyntactic violations. *Language and cognitive processes,* 13(1), 21-58.

Davis, M. H., & Gaskell, M. G. (2009). A complementary systems account of word learning: neural and behavioural evidence. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *364*(1536), 3773–800.

Dekeyser, R. (2012). Interactions Between Individual Differences, Treatments, and Structures in SLA. *Language Learning*, *62*(SUPPL. 2), 189–200.

DeLong, K. a, Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience*, *8*(8), 1117–1121.

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21.

Dronkers, N. F., Wilkins, D. P., Van Valin, R. D., Redfern, B. B., & Jaeger, J. J. (2004). Lesion analysis of the brain areas involved in language comprehension. *Cognition*, *92*(1-2), 145–177.

Ettlinger, M., Morgan-Short, K., Faretta-Stutenberg, M., & Wong, P. C. M. (2015). The Relationship Between Artificial and Second Language Learning. *Cognitive Science*, n/a–n/a.

Federmeier, K. D., & Kutas, M. (1999). A Rose by Any Other Name : Long-Term Memory Structure and Sentence Processing. *Journal of Memory and Language*, *41*(December), 469–495.

Federmeier, K. D., McLennan, D. B., De Ochoa, E., & Kutas, M. (2002). The impact of semantic memory organization and sentence context information on spoken language processing by younger and older adults: an ERP study. *Psychophysiology*, *39*(2), 133–146.

Federmeier, K. D., Wlotko, E. W., De Ochoa-Dewald, E., & Kutas, M. (2007). Multiple effects of sentential constraint on word processing. *Brain Research*, *1146*(1), 75–84.

Finn, A. S., Hudson Kam, C. L., Ettlinger, M., Vytlacil, J., & D'Esposito, M. (2013). Learning language with the wrong neural scaffolding: the cost of neural commitment to sounds. *Frontiers in Systems Neuroscience*, *7*(November), 85.

Flöel, A., de Vries, M. H., Scholz, J., Breitenstein, C., & Johansen-Berg, H. (2009). White matter integrity in the vicinity of Broca's area predicts grammar learning success. *NeuroImage*, *47*(4), 1974–81.

Friederici, A. D., Hahne, A., Mecklinger, A. (1996). Temporal structure of syntactic parsing: early and late event-related brain potential effects. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 22(5):1219–1248.

Friederici, A. D., Hahne, A., & Saddy, D. (2002). Distinct neurophysiological patterns reflecting aspects of syntactic complexity and syntactic repair. *Journal of Psycholinguistic Research*, *31*(1), 45–63.

Friederici, A. D., & Frisch, S. (2000). Verb argument structure processing: The role of verb-specific and argument-specific information. *Journal of Memory and language*, *43*(3), 476-507.

Friederici, A. D., Steinhauer, K., & Pfeifer, E. (2002). Brain signatures of artificial language processing: Evidence challenging the critical period hypothesis. *Proceedings of the National Academy of Sciences*, *99*(1), 529-534.

Golestani, N., Molko, N., Dehaene, S., LeBihan, D., & Pallier, C. (2007). Brain structure predicts the learning of foreign speech sounds. *Cerebral Cortex*, *17*(3), 575–582.

Golinkoff, R. M., & Hirsh-Pasek, K. (2008). How toddlers begin to learn verbs. *Trends in Cognitive Sciences*, *12*(10), 397–403.

Greenhouse, S., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, *24*(2), 95–112.

Groppe, D. M., Urbach, T. P., & Kutas, M. (2011). Mass univariate analysis of event-related brain potentials/fields I: A critical tutorial review. *Psychophysiology*, *48*(12), 1711–1725.

Gunter, T. C., Stowe, L. A., & Mulder, G. (1997). When syntax meets semantics. *Psychophysiology*, *34*(6), 660-676.

Hagoort, P., Brown, C., & Groothusen, J. (1993). The syntactic positive shift (sps) as an erp measure of syntactic processing. *Language and*

Cognitive Processes, *8*(4), 439–483.

Hagoort, P., & Indefrey, P. (2014). The Neurobiology of Language Beyond Single Words. *Annual Review of Neuroscience*, *37*, 347–362.

Hagoort, P., & Brown, C. M. (2000). ERP effects of listening to speech compared to reading: the P600/SPS to syntactic violations in spoken sentences and rapid serial visual presentation. *Neuropsychologia*, *38*(11), 1531-1549.

Hahne, A., Mueller, J. L., & Clahsen, H. (2006). Morphological processing in a second language: behavioral and event-related brain potential evidence for storage and decomposition. *Journal of Cognitive Neuroscience*, *18*(1), 121–34.

Hakuta, K., Bialystok, E., & Wiley, E. (2003). Critical evidence a test of the critical-period hypothesis for second-language acquisition. Psychological Science, 14(1), 31-38.

Hestvik, A., Bradley, E., & Bradley, C. (2012). Working Memory Effects of Gap-Predictions in Normal Adults: An Event-Related Potentials Study. *Journal of Psycholinguistic Research*, *41*(6), 425–438.

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews. Neuroscience*, *8*(5), 393–402.

Hudson Kam, C. L., & Chang, A. (2009). Investigating the cause of language regularization in adults: Memory constraints or learning effects? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(3), 815–821.

Hudson Kam, C. L., & Newport, E. (2005). Regularizing Unpredictable Variation: The Roles of Adult and Child Learners in Language Formation and Change. *Language Learning and Development*, *1*(2), 151–195.

Hudson Kam, C. L., & Newport, E. L. (2009). Getting it right by getting it wrong: When learners change languages. *Cognitive Psychology*, *59*(1), 30–66.

Hultén, A., Vihla, M., Laine, M., & Salmelin, R. (2009). Accessing newly learned names and meanings in the native language. *Human Brain Mapping*, *30*(3), 976–89.

Inoue, K., & Osterhout, L. (2010). *Sentence processing as a neural seesaw*.

Itzhak, I., Pauker, E., Drury, J. E., Baum, S. R., & Steinhauer, K. (2010). Event-related potentials show online influence of lexical biases on prosodic processing. *Neuroreport*, *21*(1), 8–13.

Jung, T.-P., Makeig, S., Lee, T., Mckeown, M. J., Brown, G., Bell, A., & Sejnowski, T. (2000). Independent component analysis of biomedical signals. *Proceedings If the 2nd International Workshop on Independent Component Analysis and Blind Source Separation*, (1), 633–644.

Kaan, E., Harris, A., Gibson, E., & Holcomb, P. (2000). The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes*, *15*(2), 159–201.

Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman brief intelligence test* (2nd ed.). Circle Pines, MN: AGS Publishing.

Kaufman, S. B., DeYoung, C. G., Gray, J. R., Brown, J., & Mackintosh, N. (2009). Associative learning predicts intelligence above and beyond working memory and processing speed. Intelligence, 37(4), 374-382.

Khalifian, N., Stites, M. C., & Laszlo, S. (2015). Relationships between event-related potentials and behavioral and scholastic measures of reading ability: A large-scale, cross-sectional study. *Developmental Science*, n/a–n/a.

Kos, M., van den Brink, D., & Hagoort, P. (2012). Individual Variation In The Late Positive Complex To Semantic Anomalies. *Frontiers in Psychology*, *3*(SEP), 1–10.

Kutas, M. (1993). In the company of other words: Electrophysiological evidence for single-word and sentence context effects. *Language and Cognitive Processes*.

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*, *62*(August), 621–647.

Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature*, *307*(5947), 161–163.

Landi, N., & Perfetti, C. a. (2007). An electrophysiological investigation of semantic and phonological processing in skilled and less-skilled comprehenders. *Brain and Language*, *102*(1), 30–45.

Laszlo, S., & Federmeier, K. D. (2009). A Beautiful Day in the Neighborhood: An Event-Related Potential Study of Lexical Relationships and Prediction in Context. *Journal of Memory and Language*, *61*(3), 326–338.

Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: (de)constructing the N400. *Nature Reviews. Neuroscience*, *9*(12), 920–933.

Lesaux, N., Koda, K., Siegel, L., & Shanahan, T. (2006). Development of literacy. In D. August & T. Shanahan (Eds.), *Developing literacy in second language learners* (pp. 75–122). Mahwah, NJ: Erlbaum.

Li, P., & Grant, A. (2015). Second language learning success revealed by brain networks. *Bilingualism: Language and Cognition*, (June), 1–8.

Linck, J. a., Hughes, M. M., Campbell, S. G., Silbert, N. H., Tare, M., Jackson, S. R., … Doughty, C. J. (2013). Hi-LAB: A New Measure of Aptitude for High-Level Language Proficiency. *Language Learning*, *63*(3), 530–566.

López-Barroso, D., Catani, M., Ripollés, P., Dell'acqua, F., Rodríguez-Fornells, A., & de Diego-Balaguer, R. (2013). Word learning is mediated by the left arcuate fasciculus. *Proceedings of the National Academy of Sciences of the United States of America*.

Lopez-Calderon, J., & Luck, S. J. (2014). ERPLAB: an open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience*, *8*(April), 1–14.

Loui, P., Li, H. C., & Schlaug, G. (2011). White matter integrity in right hemisphere predicts pitch-related grammar learning. *NeuroImage*, *55*(2), 500–7.

Maindonald, J., & Braun, W. J. (2014). DAAG: Data Analysis And Graphics data and functions. R package version 1.20. http://CRAN.R-project.org/package=DAAG

Mclaughlin, J., & Osterhout…, L. (2004). Neural correlates of second-language word learning: minimal instruction produces rapid change. *Nature Neuroscience*.

McLaughlin, J., Tanner, D., Pitkänen, I., Frenck□Mestre, C., Inoue, K., Valentine, G., & Osterhout, L. (2010). Brain Potentials Reveal Discrete Stages of L2 Grammatical Learning. *Language Learning*, *60*, 123–150.

Mclaughlin, J., Tanner, D., Pitkänen, I., Frenck□Mestre, C., Inoue, K.,

Valentine, G., & Osterhout, L. (2010). Brain Potentials Reveal Discrete Stages of L2 Grammatical Learning. *Language Learning*, *60*, 123–150.

Mecklinger, A., Schriefers, H., Steinhauer, K., & Friederici, A. D. (1995). Processing relative clauses varying on syntactic and semantic dimensions: An analysis with event-related potentials. *Memory & Cognition*, *23*(4), 477-494.

Melby-Lervåg, M., & Lervåg, A. (2011). Cross-linguistic transfer of oral language, decoding, phonological awareness and reading comprehension: A meta-analysis of the correlational evidence. *Journal of Research in Reading*, *34*(1), 114–135.

Mestres-Misse, A., Rodriguez-Fornells, A., & Munte, T. (2006). Watching the Brain during Meaning Acquisition. *Cerebral Cortex*, *17*(8), 1858–1866.

Midgley, K. J., Holcomb, P. J., & Grainger, J. (2009). Language effects in second language learners and proficient bilinguals investigated with event-related potentials. *Journal of Neurolinguistics*, *22*(3), 281–300.

Molinaro, N., Barber, H. a., & Carreiras, M. (2011). Grammatical agreement processing in reading: ERP findings and future directions. *Cortex*, *47*(8), 908–930.

Morgan-Short, K., Finger, I., Grey, S., & Ullman, M. T. (2012a). Second language processing shows increased native-like neural responses after months of no exposure. *PloS One*, *7*(3), e32974.

Morgan-Short, K., Steinhauer, K., Sanz, C., & Ullman, M. T. (2012b). Explicit and implicit second language training differentially affect the achievement of native-like brain activation patterns. *Journal of cognitive neuroscience*, *24*(4), 933-947.

Münte, T. F., Heinze, H. J., & Mangun, G. R. (1993). Dissociation of brain activity related to syntactic and semantic aspects of language. *Journal of Cognitive Neuroscience*, *5*(3), 335-344.

Münte, T. F., Matzke, M., & Johannes, S. (1997). Brain activity associated with syntactic incongruencies in words and pseudo-words. *Journal of Cognitive Neuroscience*, *9*(3), 318-329.

Nakano, H., Saron, C., & Swaab, T. Y. (2010). Speech and span: working memory capacity impacts the use of animacy but not of world knowledge during spoken sentence comprehension. *Journal of Cognitive Neuroscience*, *22*(12), 2886–2898.

Newman, A. J., Tremblay, A., Nichols, E. S., Neville, H. J., & Ullman, M. T. (2012). The Influence of Language Proficiency on Lexical Semantic Processing in Native and Late Learners of English. *Journal of Cognitive Neuroscience*, *24*(5), 1205–1223.

Newport, E. L. (1990). Maturational Constraints on Language Learning. *Cognitive Science*, *14*, 11–28.

Nyikos, M. (1990). Sex‐Related Differences in Adult Language Learning: Socialization and Memory Factors. The Modern Language Journal, 74(3), 273-287.

Osterhout, L. (1997). On the brain response to syntactic anomalies: manipulations of word position and word class reveal individual differences. *Brain and Language*, *59*(3), 494–522.

Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, *31*(6), 785–806.

Osterhout, L., & Holcomb, P. J. (1993). Event-related potentials and syntactic anomaly: Evidence of anomaly detection during the perception of continuous speech. *Language and Cognitive Processes*, *8*(4), 413–437.

Osterhout, L., Holcomb, P. J., & Swinney, D. A. (1994). Brain potentials elicited by garden-path sentences: evidence of the application of verb information during parsing. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *20*(4), 786–803.

Osterhout, L., Kim, A., & Kuperberg, G. (2012). The Neurobiology of Sentence Comprehension. In M. J. Spivey, M. F. Joanisse, & K. McRae (Eds.), *Cambridge Handbook of Psycholinguistics* (pp. 365–390). Cambridge University Press.

Osterhout, L., McLaughlin, J., Pitkänen, I., Frenck-Mestre, C., & Molinaro, N. (2006). Novice learners, longitudinal designs, and event-related potentials: A means for exploring the neurocognition of second language processing. *Language Learning*, *56*(SUPPL. 1), 199–230.

Osterhout, L. & Hagoort, P. (1999). A superficial resemblance does not necessarily mean you are part of the family: Counterarguments to Coulson, King and Kutas (1998) in the P600/SPS-P300 debate. *Language and Cognitive Processes*,*14*(1), 1-14.

Osterhout, L., McKinnon, R., Bersick, M., & Corey, V. (1996). On the language specificity of the brain response to syntactic anomalies: Is the syntactic positive shift a member of the P300 family?. *Journal of Cognitive Neuroscience*, *8*(6), 507-526.

Osterhout, L., Kim, A., & Kuperberg, G. R. (2012). The neurobiology of sentence comprehension. *The Cambridge handbook of psycholinguistics*, 365-389.

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*(12), 2728–44.

Patel, A. D., Gibson, E., Ratner, J., Besson, M., & Holcomb, P. J. (1998). Processing syntactic relations in language and music: an event-related potential study. *Journal of Cognitive Neuroscience*, *10*(6), 717–733.

Perfetti, C. A., Wlotko, E. W., & Hart, L. A. (2005). Word learning and individual differences in word learning reflected in event-related potentials. *Journal of Experimental Psychology: Learning Memory and Cognition*, *31*(6), 1281–1292.

Pinker, S. (1984). *Language learnability and language development*. Cambridge, MA: Harvard University Press.

Prat, C.S., Yamasaki, B.L., Kluender, R.A., Stocco, A. (2016). Resting-state qEEG predicts rate of second language learning in adults. *Brain and Language*, 157:44–50.

Qi, Z., Han, M., Garel, K., Chen, E. S., & Gabrieli, J. D. E. (2015). White-matter structure in the right hemisphere predicts Mandarin Chinese learning success. *Journal of Neurolinguistics*, *33*, 14–28.

R Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing,Vienna, Austria. URL http://www.R-project.org/.

Rosseel Y. (2012) lavaan: An R Package for Structural Equation Modeling. *J Stat Softw*. 48(2):1–36. doi:10.18637/jss.v048.i02.

Sakai, K. L. (2005). Language Acquisition and Brain Development. *Science*, *310*(5749), 815–819.

Salisbury, D. F. (2004). Semantic memory and verbal working memory correlates of N400 to subordinate homographs. *Brain and Cognition*, *55*(2), 396–399.

Sassenhagen, J., Schlesewsky, M., & Bornkessel-Schlesewsky, I. (2014). The P600-as-P3 hypothesis revisited: single-trial analyses reveal that the late EEG positivity following linguistically deviant material is reaction time aligned.*Brain and language*, *137*, 29-39.

Service, E., Helenius, P., Maury, S., & Salmelin, R. (2007). Localization of syntactic and semantic brain responses using magnetoencephalography. *Journal of Cognitive Neuroscience*, *19*(7), 1193–1205.

Snow, C. E., & Hoefnagel-Hohle, M. (1978). The Critical Period for Language Acquisition: Evidence from Second Language Learning. *Child Development*, *49*(4), 1114.

Sparks, R. L. (2012). Individual Differences in L2 Learning and Long-Term L1-L2 Relationships. *Language Learning*, *62*(SUPPL. 2), 5–27.

Sparks, R. L., Patton, J., & Ganschow, L. (2012). Profiles of more and less successful L2 learners: A cluster analysis study. *Learning and Individual Differences*, *22*(4), 463–472.

Sparks, R., Patton, J., Ganschow, L., & Humbach, N. (2009). Long-term crosslinguistic transfer of skills from L1 to L2. *Language Learning*, *59*(1), 203–243.

Steiger, J.H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin*. 87:245–251.

Steinhauer, K., & Drury, J. E. (2012). On the early left-anterior negativity (ELAN) in syntax studies. *Brain and language*, *120*(2), 135-162.

Tan, L. H., Chen, L., Yip, V., Chan, A. H. D., Yang, J., Gao, J.-H., & Siok, W. T. (2011). Activity levels in the left hemisphere caudate-fusiform circuit predict how well a second language will be learned. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(6), 2540–4.

Tanner, D., & Van Hell, J. G. (2014). ERPs reveal individual differences in morphosyntactic processing. *Neuropsychologia*, 1–13.

Tanner, D., Inoue, K., Osterhout, L. (2014). Brain-based individual differences in online L2 grammatical comprehension. *Bilingualism: Language and Cognition*, 17:277–293.

Tanner, D., Morgan□Short, K., & Luck, S. J. (2015). How inappropriate high□pass filters can produce artifactual effects and incorrect conclusions in ERP studies of language and cognition. *Psychophysiology*, *52*(8), 997-1009.

Tse, C., Lee, C., Sullivan, J., Garnsey, S., Dell, G., Fabiani, M., & Gratton, G. (2007). Imaging cortical dynamics of language processing with the event-related optical signal. *Proceedings of the National Academy of Sciences*, *104*(43), 17157.

Ullman, M. T., Corkin, S., Coppola, M., Hickok, G., Growdon, J. H., Koroshetz, W. J., & Pinker, S. (1997). A Neural Dissociation within Language: Evidence that the Mental Dictionary Is Part of Declarative Memory, and that Grammatical Rules Are Processed by the Procedural System. *Journal of Cognitive Neuroscience*, *9*(2), 266–276.

Van Petten, C., & Kutas, M. (1990). Interactions between sentence context and word frequency in event-related brain potentials. *Memory & Cognition*, *18*(4), 380–393.

Van Petten, C., Weckerly, J., McIsaac, H. K., & Kutas, M. (1997). Working memory capacity dissociated lexical and sentental context effects. *Psychological Science*, *8*(3), 238–242.

Ventura-Campos, N., Sanjuán, A., González, J., Palomar-García, M.-Á., Rodríguez-Pujadas, A., Sebastián-Gallés, N., … Ávila, C. (2013). Spontaneous brain activity predicts learning ability of foreign sounds. *The Journal of Neuroscience*, *33*(22), 9295–305.

Veroude, K., Norris, D. G., & Shumskaya, E. (2010). Functional connectivity between brain regions involved in learning words of a new language. *Brain and Language*, *113*(1), 21–27.

Vignolo, L. (1988). The anatomical and pathological basis of aphasia. In F. C. Rose, R. Whurr, & M. A. Wyke (Eds.), *Aphasia*. Amsterdam: North Holland.

Vos, S. H., & Friederici, A. D. (2003). Intersentential syntactic context effects on comprehension: the role of working memory. *Cognitive Brain Research*, *16*(1), 111–122.

White, E. J., Genesee, F., & Steinhauer, K. (2012). Brain Responses before and after Intensive Second Language Learning: Proficiency Based Changes and First Language Background Effects in Adult Learners. *PLoS ONE*, *7*(12).

Wong, F. C. K., Chandrasekaran, B., Garibaldi, K., & Wong, P. C. M. (2011). White Matter Anisotropy in the Ventral Language Pathway Predicts Sound-to-Word Learning Success. *Journal of Neuroscience*, *31*(24), 8780–8785.

Wong, P. C. M., Perrachione, T. K., & Parrish, T. B. (2007). Neural characteristics of successful and less successful speech and word learning in adults. *Human Brain Mapping*, *28*(10), 995–1006.

Yum, Y. N., Midgley, K. J., Holcomb, P. J., & Grainger, J. (2014). An ERP study on initial second language vocabulary learning. *Psychophysiology*, 51(4), 364-373.