COLUMBIA UNDERGRADUATE SCIENCE JOURNAL

Neural Oscillations as Predictors of Variability in Second Language Proficiency

Victoria Ogunniyi[1]*, David Abugaber[2], Irene Finestrat[2], Alicia Luque[4], Kara Morgan-Short[2,3]

[1]Department of Biological Sciences, University of Illinois at Chicago, Chicago, IL
[2]Department of Hispanic Linguistics, University of Illinois at Chicago, Chicago, IL
[3]Department of Psychology, University of Illinois at Chicago, Chicago, IL
[4]Department of Language and Culture, UiT The Arctic University of Norway, Tromsø (Norway)

KEYWORDS: Second language acquisition, language proficiency, quantitative EEG, psycholinguistics, resting-state studies

ABSTRACT: Understanding what traits facilitate second language (L2) learning has been the focus of many psycholinguistic studies for the last thirty years. One source of insight comes from quantitative electroencephalography (qEEG), i.e., electrical brain activity recorded from the scalp. Using qEEG, [1] found that functional brain connectivity is predictive of language learning ability. This study extends Prat et al. in investigating the association of qEEG measures for two measures of L2 proficiency, namely: 1. a grammaticality judgment task, wherein participants read and identified Spanish sentences as either correct or incorrect based on possible grammar violations, and 2. a standardized Spanish proficiency test (DELE). Participants were low-intermediate L2 learners recruited from third- and fourth-semester university Spanish classes. Spectral power and coherence within and across six different regions were analyzed for correlations with either scores on the grammaticality judgment task or on the DELE. Follow-up linear regression models based on significant qEEG correlates explained up to 11% of variance in DELE scores but none of the variance in grammaticality judgment task performance. Negative correlations were found between theta frequency coherence and the DELE. Because theta activity has been associated with episodic and working memory performance, these findings suggest that less proficient learners might utilize memory-based strategies more often to compensate for their lack of familiarity with the L2.

INTRODUCTION

Understanding what characteristics underlie successful learning is not only pedagogically important but also intriguing from a cognitive standpoint. In the study of linguistics, this question has frequently been examined in regard to second language (L2) acquisition (for a review, see [2]). Researchers have used a variety of approaches to investigate which cognitive abilities are correlated with higher L2 proficiency. Going beyond behavioral measures for psycholinguistic and cognitive constructs

[2], neuroimaging techniques such as electroencephalography (EEG) allow us to directly quantify and connect brain activity to the processes involved in language processing [3].

This study addresses the question of whether EEG can explain variance on outcomes of L2 proficiency. We first review the literature on qEEG and its use as a neurocognitive measure in L2 studies, focusing on the studies that examine resting-state qEEG as a potential factor in language learning ability and language proficiency. Then, in a conceptual replication and extension of [1], we describe the results of our study in which EEG measures of resting-state brain rhythms were explored for relationships with behavioral measures of L2 proficiency. The potential neurocognitive and pedagogical implications of these findings are expounded on in the discussion section.

Through our conceptual replication of [1], which focused on measures of L2 learning, we aimed to examine whether prior research on resting-state qEEG and L2 learning rate would extend to the construct of L2 proficiency, that is, whether a learner's intrinsic pattern of brain rhythms is associated with their L2 abilities.

Background on qEEG

EEG is an electrophysiological technique that utilizes electrodes placed on the scalp to measure changes in voltage between electrodes [4, 5]. These transient shifts in electric potential are caused by the electrical activity of neurons. Due to their proximity to the scalp, pyramidal neurons, which project information to neurons in local regions, produce most of the postsynaptic potentials recorded by EEG [5]. Although EEG data cannot attribute the electrical activity to specific brain regions, its temporal resolution allows researchers to track changes in brain activity to the millisecond. Thus, since the 1960s, EEG has been a widely used tool in cognitive studies [6].

There are various methods that can be used to analyze EEG data. The analysis of raw EEG, or qEEG, data yields useful information via neural oscillations, that is, rhythmic or repetitive patterns of neural activity. In contrast to more common methods that analyze voltage amplitudes within given time windows tied to a stimulus, such as event related potentials, qEEG has the advantage of providing information about neural activity occurring before and after the onset of a stimulus, or even in the absence of any particular stimulus. In prior literature, qEEG has been used in disparate domains, such as serving as evidence of mental dysfunction in criminal cases [7], providing neurofeedback for therapy patients with ADHD [8], and predicting learners' aptitude for learning computer programming languages [9].

Neural oscillations can be quantified through three measures: synchrony, amplitude, and coherence. The first measure, synchrony, describes whether neural oscillations are increasing or diminishing during a cognitive process [6]. By measuring the phase synchronization and desynchronization of the neural oscillations in this way, researchers can map larger interactions among the brain's networks and demonstrate patterns of activation. This may provide insight into a possible solution to the binding problem, which asks how the brain integrates separate streams of information into one cohesive mental representation [5]. The second measure, amplitude, describes local changes in synchrony. Amplitude within a certain frequency band is also often referred to as power. Though an increase in power does not always reflect the presence of oscillations, sustained power increases within a narrow frequency range is usually a good indicator that oscillations are likely present at that frequency [6]. The third measure, coherence, describes the similarity in waveform properties and the stability of phase differences between two oscillations across brain regions [6]. Though these three measures do not encompass all possible properties of oscillations, they can describe how oscillations represent activation and suppression of different neural networks, how wave amplitude is related to increased general activation, and how oscillations communicate over long distances.

Though their specific ranges vary between studies, there are five primary frequency bands that neural oscillations can be divided into [5, 6, 10]: delta (1-2 Hz), theta (3-7 Hz), alpha (8-12 Hz), beta (13-30 Hz), and gamma (30-200 Hz). These bands have been implicated in a variety of cognitive processes. For instance, prior research has suggested a relationship between oscillation frequency and the range of neural network interactions: the lower frequencies, alpha and below, represent local interactions, whereas higher frequencies represent interactions between more distant brain regions [10].

In regard to the more general applications of qEEG, all of the frequency bands have been shown to play a role in memory. For instance, the delta band, which is the predominant frequency found in slow wave sleep, has been associated with memory consolidation [5]. In particular, delta oscillations facilitate the formation of declarative memory, the memory of one's experiences and explicit knowledge. Similarly, theta oscillations have been implicated in both working memory and long-term memory retrieval [5, 11]. Theta oscillations have primarily been observed in the cortex, further echoing these memory functions [10, 11]. Alpha oscillations, which are the most prominent in the adult brain, are related to attention paid to external stimuli [5, 6]. More importantly, alpha plays a role in blocking irrelevant information in working memory. Additionally, alpha desynchronization and reductions in alpha power result in more successful information encoding [10]. Recent studies have shown that the beta band, which is mostly generated in the fronto-central region of the brain, also plays a role in regulating information stored in working memory [12]. The gamma band, which has been observed in the cortex [5], has been associated with short- and long-term memory maintenance [6]. Additionally, increases in gamma activity is anti-correlated with beta activity levels [12].

Additionally, the bands have been shown to play significant roles in stimulus-based language tasks. For instance, increased power in the theta band, which occurs during grammatical violations and sentence contexts that are difficult to interpret, reflects its involvement in lexical-semantic memory retrieval. Similarly, the alpha band helps organize information stored in short-term memory during sentence comprehension [10]. Regarding the higher frequency bands, gamma and beta have been associated with unifying related word meanings and similar grammatical forms, respectively. More specifically, the gamma band has been attributed to semantic unification [6, 10], which is supported by the observed decreases in gamma power in response to phrases with unclear meanings and idiomatic expressions [10]. Conversely, the beta band is involved in syntactic unification [5, 6, 10]. Beta oscillations sentence to lower processing regions [10]. By reflecting both domain-general and stimulus-specific cognitive functions, qEEG has proven to be a useful neurocognitive measure in psycholinguistic studies.

The Use of qEEG in L2 Studies

Several L2 studies have tested the relationship between qEEG measures and L2 constructs. These studies have generally addressed two issues: L2 proficiency [13-16], which describes a learner's language abilities at a given point in time, and L2 grammatical learning [17-19], which describes how learners better understand the rules of a language with increasing proficiency. Regarding L2 proficiency studies, results have shown that highly proficient L2 learners differ in qEEG measures from less proficient L2 learners [14-16]. When comparing differently proficient participant groups, significant differences are found in the timing and location of oscillatory activity, especially regarding lateralization of the location of the oscillations between the right and left hemisphere. Though the delta band is not frequently analyzed in these studies, one study found no significant group differences [13]. For grammatical learning studies, in both natural and artificial grammar learning tasks, higher relative power and coherence in the higher frequency bands (>8 Hz) were associated with higher proficiency and increased in prevalence over time, whereas higher relative power and coherence in the lower frequency bands were associated with lower proficiency and decreased over time [16, 17, 19]. Gauging learning by assessing oscillatory responses to grammatical violations, other research has also found that the higher frequency bands were elicited in both semantic and syntactic conditions [17].

A relatively new approach in L2 studies is to examine whether resting-state qEEG measures (i.e., that are taken when the brain is “idling” in the absence of any explicit task, as opposed to stimulus-related qEEG measures) are associated with the rate at which an individual acquires second language abilities, or L2 learning, and proficiency. As of now, very few studies have tested the correlation between neural activity occurring in the absence of a stimulus, or resting-state qEEG, and L2 learning rates [1, 20]. In these studies, native-English speakers learned a second language over the span of a few months. Prior to this learning period, resting-state EEG was performed, and various behavioral measures were collected as additional outcome measures. In the first of these studies, [20] found that, when entered as predictors into a regression model, resting-state qEEG measures explained up to 60% of the variance observed in L2 learning rates, meaning that learning rate accounted for more than half of the variability in the data. Though the most predictive frequency range was found to be the low-beta range (13-14.5 Hz), power in the beta and gamma frequency bands recorded primarily over right hemisphere electrode regions were found to be the strongest predictors of L2 learning ability in general. The authors also found that alpha power over frontal and temporal electrodes and low-beta power over temporal regions were indicative of better language learning ability. As with other studies, greater activation in left hemisphere electrode sites was associated with lower L2 proficiency.

In a later study conducted by [1] with a higher sample size, the authors tested whether resting-state qEEG measures were significant predictors of different L2 learning measures. Similar to their earlier study, the results implicated the qEEG activity in the right hemisphere with greater L2 learning ability. Simultaneous regression analyses were run on three outcome variables: L2 learning rate, total speech attempts, and performance on a declarative memory posttest. Mean right posterior beta power was found to be a significant predictor of L2 learning rate and total speech attempts. Frontotemporal to posterior coherence in the right hemisphere was found to be a significant predictor of performance on the declarative memory posttest, whereas mean within left posterior coherence across all frequencies was a significant predictor of total speech attempts. Altogether, these results indicate that mean beta power over posterior electrode regions plays a significant role in L2 ability.

In all, due to their potential advantages over stimulus-locked measures, qEEG measures have been increasingly used in L2 studies. Though some frequency bands have been more frequently studied than others, all five of the classic bands have been implicated in various language functions in some way. Across various research designs, qEEG has been used to illustrate large-scale patterns of brain activation during language processing. However, relatively little research has explored the potential of resting-state paradigms for qEEG in L2 psycholinguistics. For example, although [1, 20] have examined whether qEEG can predict individual differences in L2 learning rate, speech attempts during learning, and L2 declarative knowledge, no study has yet examined whether or how resting-state qEEG measures may be predictive of L2 proficiency, which is the desired final outcome of L2 learning.

Purpose of Research

The goal of this study is to investigate the potential association between resting-state qEEG and L2 proficiency. The current study expands on the research design of previous resting-state L2 studies. At present, resting-state qEEG studies have involved extensive training sessions with participants in the initial stages of L2 learning. However, the participant population utilized in this study comprised L2 learners at a low-intermediate stage who were recruited from third- and fourth-semester university-level Spanish courses. In connecting the qEEG measures with observed variability in language proficiency, we reasoned that second-year Spanish learners with the qEEG profiles most conducive to effective processing (e.g., through memory functions and other related cognitive processes) would have gained the most Spanish proficiency from their classes. Additionally, this study has the advantage of analyzing a greater number of electrodes than prior research [1, 20], providing higher spatial resolution for measuring electrical activity on the scalp. As with prior resting-state studies, a variety of qEEG measures were analyzed, including spectral power and coherence. Considering the issues above, this study specifically investigated two research questions:

Research Question 1: Is mean spectral power calculated from resting-state qEEG data associated with L2 proficiency, as assessed by two Spanish proficiency tasks?

Research Question 2: Is mean within- and between-network coherence calculated from resting-state qEEG data associated with L2 proficiency, as assessed by two Spanish proficiency tasks?

Following the results of [1], we predicted that the qEEG measures that were most likely to show a relationship with the two L2 proficiency tasks were mean beta power and frontotemporal-to-posterior coherence. In spite of these tentative predictions, we sought to replicate [1]'s methods as closely as possible in reproducing their two-step exploratory analysis strategy (i.e., pairwise correlations followed by multiple regression using significant correlates as predictors) on all frequency bands. As we intended our analysis itself to be strictly exploratory rather than confirmatory, we included all frequency bands in the analysis. In order to assess L2 proficiency, we decided to administer two assessments, one that reflects specific grammatical knowledge acquired through the learners' class (the grammaticality judgment task), and one that is a more general and widely measure in the L2 literature (Diplomas de Español como Lengua Extranjera), see Methods. Given that prior research [21] has found a distinction between automatic and controlled language processing, using both a timed and untimed L2 proficiency measure would allow us to examine the application of L2 knowledge in two distinct ways.

METHODS

Participants

Forty-nine participants (29 female; 20 male; mean age = 21.54; age SD = 4.14; range = 18-38) were initially recruited for participation in a two-part EEG study in which they were tested in English and in Spanish in separate testing sessions. Participants were recruited at a large, public urban university in Chicago from third- and fourth-semester Spanish language courses, which centered on developing communicative abilities and aimed to help students obtain low-intermediate proficiency by the end of the fourth-semester course. Thirteen of these participants were recruited from the third-semester Spanish course, while twenty-two of these participants were recruited from the fourth-semester Spanish course; thirteen participants did not report their course level. The recruitment process involved advertising to these language courses, after which participants self-selected whether to participate. Regarding the racial distribution of the participant population, nineteen participants identified as White/Caucasian, fourteen participants identified as Asian, six participants identified as Black/African-American, four participants identified as multiracial, and six participants' races were unreported. All the participants reported having English as a native language (even if they were simultaneous bilinguals with early exposure to other languages) and having no Spanish exposure growing up. Additionally, per our EEG criteria, all participants were right-handed as assessed by the abridged version of the Edinburgh Handedness Inventory [22], with normal or normal-to-corrected vision. This was done to ensure that our results were not confounded by uncorrected vision or differing brain activity in those who are left-handed. None of the participants reported having psychiatric, neurological, or learning disorders. All participants gave informed consent according to the standards of the University of Illinois at Chicago institutional review board and were financially compensated for their participation based on the number of hours spent participating in the study, receiving $5 for every hour spent on completing the proficiency measures, $15 for every hour spent during the EEG recording process, and a $45 bonus for completing both sessions.

Participants were only included in the analysis if they completed the resting state EEG recording along with at least one of our Spanish proficiency measures. The final number of participants included in the current analysis is 47, with 44 of these participants having completed both Spanish proficiency measures (see Table I below). A similar EEG recording procedure was performed for an L1 English reading task (not reported here). The participant data included in this study was collected over the span of two years, from 2018 to 2020.

Table I. Participant Language Characteristics

	M (SD) [Range]
Number of native languages	1.29 (0.45) [1-2]
Number of L2s	1.44 (0.53) [1-3]
Age of acquisition Spanish (years)	14.53 (4.88) [0-29]
Self-rated Spanish listening proficiency[a]	4.88 (2.01) [1-9]
Self-rated Spanish reading proficiency	5.42 (1.72) [1-8]

Note: [a]Self-rated proficiency on 0 ('none) to 10 ('perfect) scale.

Grammaticality Judgment Task

Participants read various Spanish sentences and were asked to determine whether they followed Spanish grammatical rules. The grammaticality judgment task consisted of three experimental conditions: phrase structure, subject-verb agreement, and noun-phrase violations (see Table II for examples). The phrase structure condition, wherein word order violations were introduced into a sentence by presenting a noun instead of a verb or vice-versa, consisted of 60 correct and 60 violation sen-

Table II. Examples of stimulus sentences

Item Type	Example
Phrase Structure	Ella tiene mucho dinero/*gastar que gastar/dinero en ropa. [She has a lot of money/*spend to spend/money on clothes.]
Subject-Verb	La mujer dibuja/*-dibujan en su habitación. [The lady draws/*draw in her bedroom.]
Noun Phrase	El hombre prepara estas papas/ *papa para su esposa. [The man prepares these pota-toes/*potato for his wife.]

Note: * = violation word. Italics indicate the critical correct/violation word in each sentence.

tence frames, totaling 120 sentences overall. The subject-verb agreement condition, for which the verb ending did not agree with the plurality of the subject, consisted of 60 correct sentences with a singular subject, 60 correct sentences with a plural subject, and 120 violation sentences, totaling 240 sentences. The noun-phrase condition, for which either the singular/plural status or the grammatical gender of an article (e.g., los, "the [MASC. PLURAL]"; esta, "this [FEM. SINGULAR]") did not match the noun, consisted of 124 number violation frames, 124 gender violation frames, and 248 correct sentence frames, for a total of 496 sentences. In total, 856 sentences were used across all three conditions. The sentences ranged from 5 to 12 words in length. None of the sentences contained violations in initial or final sentence positions, so as to avoid sentence "start-up" and "wrap-up" effects in the EEG [23-24], and none of the critical words were repeated between frames. To ensure that the participants would be familiar with the vocabulary contained in the sentence frames, all the words for this task were taken from a Spanish textbook used at the university at the time that data collection began [25].

Sentences were presented one word at a time on the computer screen (see Figure 1) using E-Prime 2.0 software. Instructions for the task were read orally to the participants by the experimenter. Preceding each sentence, there was a screen that read, "Rest your eyes." After three seconds, the sentence was then visually presented one word at a time. Each word was displayed in the center of the screen for 350 ms, with a 150 ms interval of blank screen before the onset of the subsequent word. Once the entire sentence was presented, a screen followed that said, "Good/Bad?" In response to this prompt, participants pressed a keyboard button to categorize the sentence as either grammatical or ungrammatical. Participants first completed a short practice block containing 8 sentences. The stimuli sentences were then presented over 4 experimental blocks. There were three 3-minute breaks during the experiment, one at the end of each block. Another EEG recording was performed over the duration of this task but was not analyzed in the current study. Participants' responses were used to calculate a d-prime (d') score, which is a metric for signal detection that accounts for response bias by comparing how often a participant correctly identifies a signal to their false-alarm rate [26].

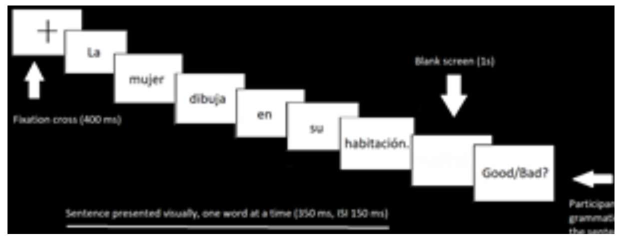

Figure 1: Diagram of a typical trial in the grammaticality judgment task.

Diplomas de Español como Lengua Extranjera (Diploma of Spanish as a Foreign Language):

A modified version of the Diplomas de Español como Lengua Extranjera (DELE, [27]) was used to assess Spanish proficiency. The three-part test was completed on a computer in the laboratory through a Qualtrics survey form. Participants were asked to read through the DELE questions and answer them at their own pace. In the first section, participants were required to read through a passage in Spanish and answer 20 fill-in-the-blank questions. Each question had 3 possible answer choices. In the second section, participants were given 10 sentences and asked to choose the answer choice that best defines the bolded word in the sentence. Each sentence also had 3 possible answer choices. The third section consisted of 19 grammatical questions . Participants were asked to select the answer choice that fit best in the context of each of the sentences. Eight of these questions had 2 possible answer choices, and the remaining questions had 4 possible answer choices. In total, participants answered 49 questions.

Procedures

Prior to testing, all participants completed pre-testing questionnaires that verified their eligibility and provided more detailed information about their language history. This included a language background questionnaire, a test-session questionnaire, and a handedness questionnaire. The language background questionnaire assessed each participant's demographic background and language history and experience (LEAP-Q, [28]). The test-session questionnaire assessed how much sleep a participant had and whether they had taken any psychoactive substances that may affect their ability to perform the task. The handedness questionnaire was used to gauge left-/right-handedness by assessing hand preference during various activities, following the standard Edinburgh Handedness Inventory [22]. The items were read to participants, who provided their answers verbally to the experimenter. Answers were recorded in computer-based survey forms.

In replicating [1]'s procedure, we collected five minutes of eyes-closed resting-state EEG following completion of the pre-testing surveys. Participants sat in a chair inside of a sound-attenuating booth. After fitting the participants for an EEG cap and placing eye electrodes, an electrolyte solution was applied to the scalp electrodes to minimize electrical impedances. Participants were then instructed to close their eyes and remain still and awake during the recording. While recording, the lights were turned off and the door of the sound booth was closed.

The EEG data was recorded using asa™ software with an ANT Neuro waveguard™ elastic cap with 32 Ag/AgCl electrodes distributed in standard and extended 10-20 system locations. Scalp impedances were lowered to 10 kΩ or below. Scalp electrodes were referenced to the common average of all the electrodes. To detect artifacts caused by eye movements, electrodes were placed above and below the right eye and on the left and right outer canthi to record a vertical electrooculogram and a horizontal electrooculogram, respectively. Using an ANT Neuro bioamplifier system (AMP-TRF40AB Refa-8 amplifier), the EEG signal was amplified to 22 bits. The signal was also recorded in DC mode, digitized with a 512 Hz sampling rate, and filtered online using a low-pass filter with a cutoff of 138.24 Hz.

Following the resting-state EEG session, the grammaticality judgment task was implemented in the sound booth (see above), and after disassembly of EEG equipment and a short break the DELE task was performed on a computer outside of the sound booth (see above).

Analyses

The qEEG data was pre-processed using the EE-GLab toolbox [29] for MATLAB [30]. To ensure that the resting-state recording was exactly five minutes, each recording was limited to 300 seconds. Seven participants had recordings that were slightly less than 300 seconds (minimum = 283 seconds) but were still included in the analysis. Each participant's recording was divided into epochs of two-second duration, with 50% overlap across epochs. These epochs were then cleaned for artifacts (e.g., from muscle movements, eyeblinks, faulty electrodes, etc.) using the pop_autorej() function from EEGLab. Participant datasets with less than 75 seconds of epoch-free recording were omitted from the final analysis, which resulted in the loss of 6 participants (12% of the data). The mean number of samples per participant was 144.44 (S.D. = 39.73).

The pre-processed data were subsequently analyzed using a modified version of the script used in the Prat et al. study [1] (available at: https://github.com/UWCCDL/QEEG) for the R scripting language [31]. Six electrode networks were defined (Figure 2): medial frontal (electrodes FP1, FPz, FP2, and Fz), left hemisphere fronto-temporal (electrodes F7, FC5, T7, C3), right hemisphere frontotemporal (F8, FC6, T8), left hemisphere posterior (CP5, CP1, P7, P3, O1), right hemisphere posterior (Cz, CP6, CP2, C4, P4, Pz), and right hemisphere posterior occipital (Oz, O2, POz, P8).

In defining the electrode clusters, we aimed to replicate Prat et al. (2019), in which qEEG channels were collapsed into networks based on phase synchrony results from an earlier L2 qEEG study. Prat et al. 2019's network definitions are technically slightly different from ours in that they used 14 electrodes rather than a 32-electrode cap, but our network definitions were aligned as closely with theirs as possible based on visual inspection of scalp maps (and in fact having a higher spatial resolution is a point in our favor in a sense). As data-driven results in favor of our network definitions (which we left out due to space limits), we replicated [1] (2019, Table I) in that independent samples t tests found that all within-network qEEG coherence values in our networks were significantly greater than all between-network coherence values, with all independent samples t-tests at p < .001. We then extracted the qEEG measures of interest, which were power and within- and between-network coherence for each of the frequency bands: theta (4-7.5 Hz), alpha (8-12.5 Hz), beta (13-29.5 Hz), and gamma (30-40 Hz).

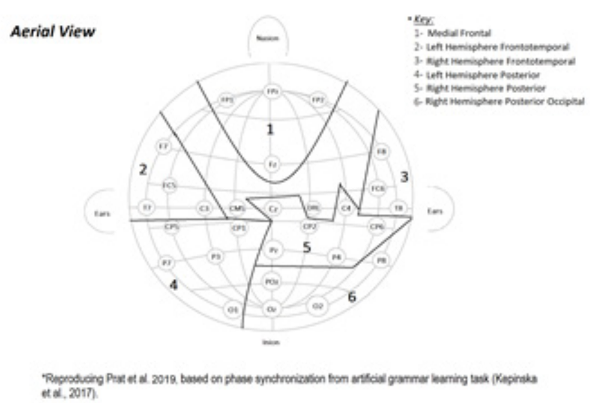

Figure 2: Diagram of the six network regions analyzed

Finally, to address the research questions, we first conducted correlations between mean power and performance on the grammaticality judgment task and the DELE, followed by correlations between mean coherence and performance on the grammaticality judgment task and the DELE. Here we report statistically significant correlations. These exploratory correlations were not corrected for multiple comparisons following the main analyses reported by [1]. For power and coherence measures that showed statistically significant correlations for either proficiency variable, we then entered them as predictors into two, separate regression analyses.

RESULTS

Individual Differences in Indicators of L2 Proficiency

Before examining the qEEG measures, descriptive statistics were examined for the two outcome measures of Spanish proficiency (see Figures 3 and 4 and Table III). With respect to the DELE, the group mean was 19 out of 49, which illustrates that the participants were overall low proficiency speakers [27]. The most proficient participant scored twice as much as the least proficient participant. With respect to the grammaticality judgment task, participants were given two scores: mean accuracy and d'. The average accuracy on the grammaticality judgment task was 76%. The average d' was 0.91. A bootstrapped simulation of chance-level d' values on 244 trials with 10,000 iterations performed using the psycho package for R [32] found a 95% confidence interval of -0.26 to 0.26. This suggests that our participants' d' values were above chance at $\alpha = 0.05$. DELE scores did seem to be above chance, as indicated by a mean accuracy of 39.5%. Considering that most of the DELE test items had 2, 3, or 4 answer choices, a minimum accuracy of 25% would at least reflect chance levels on the items with the most answer choices. DELE scores and grammaticality judgment task d' scores were not significantly correlated with one another, $r(44) = 0.27$, $p = 0.069$. As per [21], our results suggest that the grammaticality judgment task and DELE might capture somewhat different facets of L2 proficiency.

Table III. Performance on the two proficiency tasks

Performance measure	M (SD) [Range]
DELE score	19.3 (3.15) [0.13-0.26]
Grammaticality judgment task d' score	0.91 (0.74) [-0.27-2.88]
Grammaticality judgment task accuracy	0.76 (0.10) [0.51-0.98]

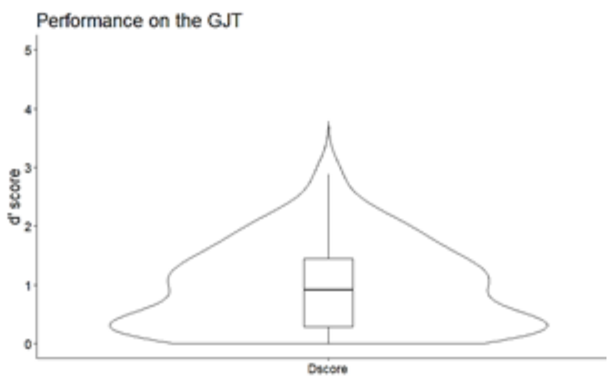

Figure 3: Participant performance on the grammaticality judgment task. The maximum possible d'score on the grammaticality judgment task was effectively 4.9.

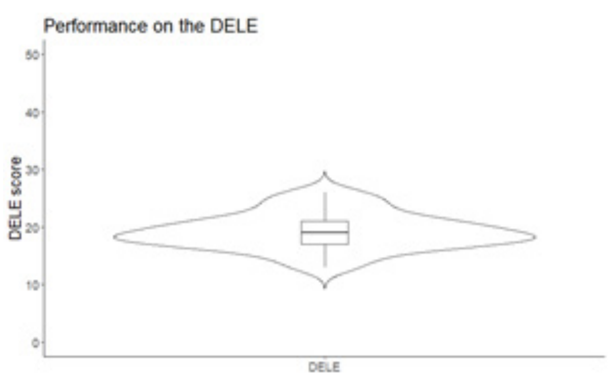

Figure 4: Participant performance on the Spanish proficiency test (DELE). The maximum possible score on the DELE was 49.

Relating Individual Differences in Resting-state qEEG Power to L2 Proficiency Variables

In order to determine the relationship between resting-state qEEG power and performance on the two proficiency tests, we performed correlation analyses between either the grammaticality judgment task or DELE scores (in separate analyses) and mean power across six electrode networks. The frequency bands of interest were theta (3-7 Hz), alpha (8-12 Hz), beta (13-30 Hz), and gamma (30-200 Hz). After conducting these analyses, none of the correlations were found to be significant. However, two positive correlations were approaching significance: medial frontal alpha power and DELE scores, $r(47) = 0.28$, $p = .052$; and left hemisphere frontotemporal alpha power and DELE scores, $r(47) = 0.27$, $p = 0.057$.

Relating Individual Differences in Resting-state qEEG Coherence to L2 Proficiency Variables

In order to determine the relationship between resting-state qEEG coherence and performance on the two proficiency tests, we performed correlation analyses between the L2 proficiency variables and mean within- and between-coherence across six electrode networks. The frequency bands of interest were theta, alpha, beta, and gamma. For the between-coherence values, each of the networks were paired together and coherence across the four frequency bands was calculated for every pair. Regarding the DELE, two significant negative correlations were found: theta coherence within the right hemisphere posterior occipital network, $r(47) = -0.31$, $p = 0.028$; theta coherence between the me-dial frontal and right hemisphere posterior occipital networks, $r(47) = -0.35$, $p = 0.012$. Regarding the grammaticality judgment task, there were no signif-icant correlations (all $p > 0.05$).

Simultaneous Linear Regression Analyses

When the two predictors of performance on the DELE (theta coherence within the right hemisphere posterior occipital network and theta coherence between medial frontal and right hemisphere posterior occipital) were entered into a simultaneous regression analysis, the overall model was found to be statistically significant, $F(2, 46) = 3.97$, $p = 0.026$, explaining up to 11% of the observed variance. However, neither theta coherence within the right hemisphere posterior occipital network ($\beta = -13.91$, $t = -1.07$, $p = 0.290$) nor theta coherence between medial frontal and right hemisphere posterior occipital ($\beta = -41.78$, $t = -1.61$, $p = 0.114$) were found to independently predict performance on the DELE (see Figures 5 and 6).

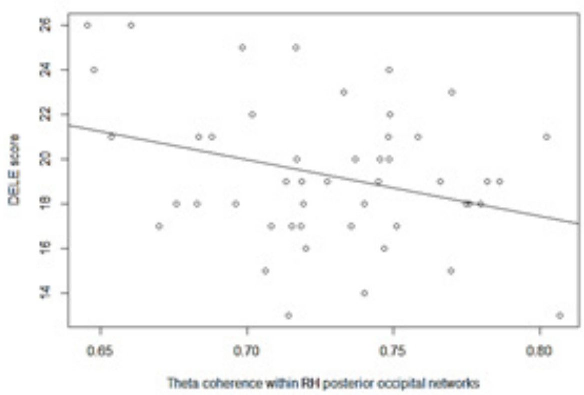

Figure 5: Regression line between the DELE and theta coherence within right hemisphere posterior occipital networks.

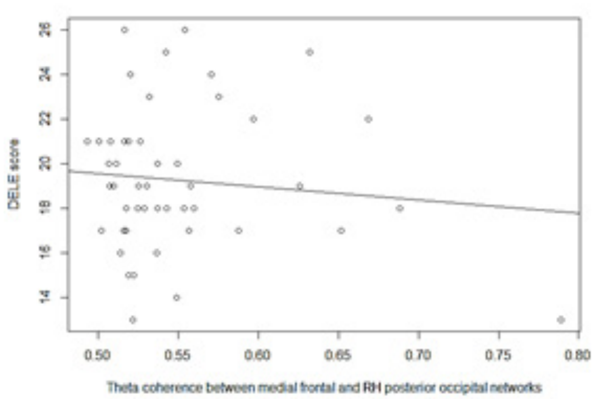

Figure 6: Regression line between the DELE and theta coherence between the medial frontal and right hemisphere posterior occipital networks.

DISCUSSION

Our results suggest that certain resting-state qEEG measures, particularly over the theta frequency band, are associated with L2 proficiency. Regarding the first research question, none of the correlations run between resting-state mean power and L2 proficiency reached significance. However, the two correlations that did approach significance were related to alpha power: the positive correlation between medial alpha power and the DELE, and the positive correlation between left hemisphere frontotemporal alpha power and the DELE. Regarding the second research question, two significant negative correlations were found between resting-state within- and between-network coherence and the DELE, one within right hemisphere posterior networks and another between medial frontal and right hemisphere posterior networks. Both significant correlations were found over the theta frequency band. After performing regressions on the significant qEEG predictors for DELE performance, the model was found to explain up to 11% of the variance. None of the variance in grammaticality judgment task performance was explained by qEEG measures.

In relating our results with those of previous studies, the correlations found between the theta frequency band and L2 proficiency were anticipated, although the direction of the relationship varied by study. In [1], theta coherence within frontal electrode regions was positively correlated with several outcome measures of L2 learning. In [16], highly proficient speakers experienced increased theta synchronization in right frontal regions during a grammar learning task. However, in the L2 proficiency study conducted by [13], lower theta coherence in frontal and occipital electrodes was observed among highly proficient speakers. In our study, the relationship between theta coherence measures and L2 proficiency was also negative. Why do we see these contradictory patterns among studies? This may be explained by the differences in language learning and language proficiency. For instance, the participant population used for language learning studies consists of speakers who were just beginning to learn an L2, whereas proficiency studies involve participants who have already had experience learning the additional language.

We understand activity in the theta frequency band to reflect several memory functions, including specifically short- and long-term memory maintenance and memory retrieval [5, 10, 33, 34]. Additionally, the theta frequency is believed to originate from the cortex [5], which plays a role in the formation of new memories. Generally, studies have found positive relationships between theta activity and learning rate in earlier stages of learning, negative relationships between theta and L2 ability in later stages [1, 16, 19]. The negative relationship found in this study may signify that more proficient participants are good at applying and retrieving grammar rules and no longer need to rely on working memory, which would result in a decreased prevalence of theta oscillations. Altogether, as suggested by [16], the negative correlations found between theta coherence and DELE performance suggest that less proficient L2 speakers have a greater reliance on memory-based strategies to compensate for their lack of familiarity with the language.

Even though the alpha frequency band was not significant, it was approaching significance, which reflects the inverse theta-alpha relationship expressed in the literature [5]. Increases in the alpha frequency have been associated with diminished attention paid to a linguistic task [10]. Interpreting the negative correlation between theta and proficiency to be the consequence of decreased working memory load, a positive correlation with alpha would suggest that more proficient participants were able to pay less attention to the task and still be successful. This may signify that the more automatic a language task is to a participant, the more likely they are to be more proficient.

The results of this study need to be considered in light of its limitations. One limitation was the proficiency measures that were employed. As mentioned in the results section, our participants did not score statistically above chance on average on the DELE. Thus, it is somewhat surprising that significant results were found for the DELE and not for the grammaticality judgment task. Perhaps the more difficult DELE test, which had been developed to test up to near-native speaker status, allowed us to detect a role related to which learners are more successful on a more challenging task. Conversely, the grammaticality judgment task was designed to reflect specific grammatical structures taught in intermediate-level Spanish courses, so we would expect participants to perform more successfully on this task overall, which they did, as evidenced by above-chance performance. More generally, because the DELE and grammaticality judgment task both reflect performance accuracy on grammatical tasks, it is important to note that they are not holistic measures. As a multidimensional construct, language proficiency encompasses all the skills necessary to engage with the language in a real-life context [2]. Thus, it would be beneficial for future research to include more time-pressured proficiency measures in future research, such as an oral elicited imitation (EIT) task. Unlike the DELE and grammaticality judgment task, which are both primarily prescriptive grammar tasks, the EIT can assess more implicit language knowledge by asking participants to listen to sentences and repeat them [35, 36]. This testing method has been found to engage long-term memory and require a higher level of language comprehension.

It is also worth noting that, while we found significant correlations with theta coherence, the majority of the qEEG measures were not associated with L2 proficiency. Though we interpreted theta to reflect the engagement of working memory, one difficulty in interpreting the results of this study, is in interpreting what cognitive processes may be reflected by the different frequency bands. Although previous research does suggest associations between activity in the frequency bands and different cognitive processes, future research will need to strengthen the validity of these claims regarding theta. In regard to the lack of a relationship between L2 proficiency and the other frequency bands, null results can be difficult to explain, but the results could be at least partially due to our processing procedures where a certain amount of the data was not included. For future analyses, we plan on implementing independent component analyses performed to correct for eye and muscle artifacts in EEG data (using ICLabel; [37]), which is expected to lead to lead to cleaner data, higher sample sizes, and improved model fits for both DELE and grammaticality judgment task.

We note three further limitations in our study that should be addressed in future research. First, future research might want to analyze data from a particular semester rather than data spanning participants from two Spanish course levels, or the course level could be included as a covariate in analyses. Second, regarding the analysis, a more precise way to define the frequency ranges of the specific frequency bands is to use individual alpha frequency (IAF) peaks. For each person, the IAF peaks at a different number, which affects the ranges of the other frequency bands [38]. Finally, given the highly exploratory nature of this study, we did not correct for multiple correlation analyses, and we entered regression predictors based on significance from the correlational analyses. Future research should conduct more conservative, confirmatory analyses on a dataset with higher statistical power to mitigate possible Type I errors. Indeed, post hoc analyses for our dataset showed that the significant correlations reported above did not survive correction for Type I error inflation using the family-wise discovery rate, which further suggests that a confirmatory study would need to be conducted to validate any of the exploratory findings reported in this study. Using this method in future research may further solidify the validity of our results or may lead to different findings.

CONCLUSION

The purpose of the current study was to investigate whether mean qEEG power and coherence are significant predictors of L2 proficiency. Based on our results, within- and between-network coherence over the theta frequency band is closely related to Spanish L2 proficiency. Because the theta frequency has been associated with memory retrieval and load, these results suggest that there is an inverse relationship between L2 proficiency and reliance on memory-based strategies for interpreting linguistic inputs. Additionally, increased theta activity may be characteristic of individuals in earlier stages of language learning. More research is needed to further validate the significance of theta in L2 proficiency, as well as to determine the importance of the other frequency bands. Ultimately, this study adds to growing literature of resting-state L2 qEEG studies, echoing the implication of intrinsic patterns of neural activity as sources of individual variation in linguistic ability. Over time, qEEG may help to reveal individual neurophysiological variations among students within a classroom, enabling educators to develop language learning strategies that will be most conducive to successful L2 outcomes for them. In other words, from the conclusions of this body of research, we might be able to identify particular cognitive processes that are associated with L2 learning and proficiency. With such information, further research could then examine how to leverage these processes in instruction.

AUTHOR INFORMATION

Corresponding Author
Victoria Ogunniyi
*vogunn2@uic.edu

Author Contributions
Victoria Ogunniyi assisted with data collection and performed the analyses along with Irene Martinez and David Abugaber. Victoria Ogunniyi wrote the manuscript. Kara

Morgan-Short, David Abugaber, Irene Finestrat, and Alicia Luque reviewed the manuscript and devised the experimental plan.

Competing Interests

The authors declare no competing financial and non-financial interests.

ACKNOWLEDGMENTS

Thank you to Kinsey Bice for her help with performing the qEEG analyses. Thank you also to all undergraduate research assistants affiliated with the Cognition of Second Language Acquisition laboratory who contributed towards data collection for this project, including Hannah Chaddah, Sarah Hassan, Gayatri Chavan, Will Camacho, Ana Hernandez, Anahi Gaytan, Krishna Bhatia, Amber Lewis, Mallory Webber, Evelyn Delgado, Daniel Henreckson, Yasiel Lacalle, and Nethaum Mizyed. We would also like to thank Sarah Malone, Christopher Powell, Crystal Galvan, and Allen Bryson of the University of Illinois at Chicago's Summer Research Opportunities Program for Undergraduates for their support in the development of this manuscript.

ABBREVIATIONS

DELE- Diplomas de Español como Lengua Extranjera
EEG – Electroencephalography
L2 – Second Language
qEEG – Quantitative Electroencephalography

REFERENCES

[1] C.S. Prat, B.L. Yamasaki, E.R. Peterson, Individual Differences in Resting-state Brain Rhythms Uniquely Predict Second Language Learning Rate and Willingness to Communicate in Adults. J Cogn Neurosci 31(1), 78-94 (2019).

[2] P. Leclercq, A. Edmonds, H. Hilton, Eds. Measuring L2 proficiency: Perspectives from SLA (Multilingual Matters, 2014).

[3] C. J. Doughty, M.H. Long, "The scope of inquiry and goals of SLA" in The handbook of second language acquisition (John Wiley & Sons, 2003), pp. 3-16.

[4] S. J. Luck, An introduction to the event-related potential technique (MIT press, 2014).

[5] K. Majumdar, A brief survey of quantitative EEG, 2nd ed. 2018 (CRC Press, 2018).

[6] M. Bastiaansen, A. Mazaheri, O. Jensen, "Beyond ERPs: Oscillatory neuronal dynamics" in The Oxford Handbook of event-related potential components, S. Luck & E. Kappenman, Eds. (Oxford University Press, 2012), pp. 31-50.

[7] G. Gluck, QEEG accepted in death penalty trial in Florida v. Nelson. Biofeedback 39(2), 74-77 (2011).

[8] M. Arns, W. Drinkenburg, J. Leon Kenemans, The effects of QEEG-informed neurofeedback in ADHD: an open-label pilot study. Appl Psycho-physiol Biofeedback 37(3), 171-180 (2012).

[9] C.S. Prat, T.M. Madhyastha, M.J. Mottarella, C.H. Kuo, Relating Natural Language Aptitude to Individual Differences in Learning Programming Languages. Sci Rep 10(1), 1-10 (2020).

[10] Y. Prystauka, A.G. Lewis, The power of neural oscillations to inform sentence comprehension: a linguistic perspective. Lang Linguist Compass 13(9), e12347 (2019).

[11] L. Meyer, The neural oscillations of speech processing and language comprehension: state of the art and emerging mechanisms. Eur J Neurosci 48(7), 2609-2621 (2018).

[12] M. Lundqvist, P. Herman, M.R. Warden,

S.L. Brincat, E.K. Miller, Gamma and beta bursts during working memory readout suggest roles in its volitional control. Nat Commun 9(1), 1-12 (2018).

[13] S. Reiterer, C. Hemmelmann, P. Rappels-berger, M.L. Berger, Characteristic functional networks in high- versus low-proficiency second language speakers detected also during native language processing: an explorative EEG coherence study in 6 frequency bands. Brain Res Cogn Brain Res 25(2), 566-578 (2005).

[14] S. Reiterer, E. Pereda, J. Bhattacharya, Measuring second language proficiency with EEG synchronization: how functional cortical networks and hemispheric involvement differ as a function of proficiency level in second language speakers. Second Lang Res 25(1), 77-106 (2009).

[15] S. Reiterer, E. Pereda, J. Bhattacharya, On a Possible Relationship between Linguistic Expertise and EEG Gamma Band Phase Synchrony. Front Psychol 2, 334 (2011).

[16] O. Kepinska, E. Pereda, J. Caspers, N.O. Schiller, Neural oscillatory mechanisms during novel grammar learning underlying language analytical abilities. Brain Lang 175, 99-110 (2017).

[17] A. Kielar, J.A. Meltzer, S. Moreno, C. Alain, E. Bialystok, Oscillatory responses to semantic and syntactic violations. J Cogn Neurosci 26(12), 2840-2862 (2014).

[18] A.G. Lewis, K. Lemhöfer, J.M. Schoffelen, H. Schriefers, Gender agreement violations modulate beta oscillatory dynamics during sentence comprehension: A comparison of second language learners and native speakers. Neuropsychologia 89, 254-272 (2016).

[19] Z.R. Cross, A. W. Corcoran, M. Schlesewsky, M.J. Kohler, I. Bornkessel-Schlesewsky, Oscillatory and aperiodic neural activity jointly predict grammar learning. BioRxiv (2020).

[20] C.S. Prat, B.L. Yamasaki, R.A. Kluender, A. Stocco, Resting-state qEEG predicts rate of second language learning in adults. Brain Lang 157, 44-50 (2016).

[21] A. Godfroid, S. Loewen, S. Jung, J. H. Park, S. Gass, R. Ellis, Timed and untimed grammaticality judgments measure distinct types of knowledge: Evidence from eye-movement patterns. Studies in Second Language Acquisition 37(2), 269-297 (2015).

[22] R.C. Oldfield, The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia 9(1), 97-113 (1971). Available at: https://www.psytoolkit.org/survey-library/handedness-ehi.html.

[23] V. Kuperman, M. Dambacher, A. Nuthmann, R. Kliegl, The effect of word position on eye-movements in sentence and paragraph reading. Q J Exp Psychol (Hove) 63(9), 1838-1857 (2010).

[24] L.A. Stowe, E. Kaan, L. Sabourin, R.C. Taylor, The sentence wrap-up dogma. Cognition 176, 232-247 (2018).

[25] J.A. Blanco, P.R. Donley, Vistas: Introduccion a la lengua espaňola, 5th ed. 2016 (Vista Higher Learning, Inc., 2016).

[26] N.A. Macmillan, C.D. Creelman, Detection theory: A user's guide (Psychology press, 2004).

[27] S. Montrul, Second language acquisition and first language loss in adult early bilinguals: exploring some differences and similarities. Second Lang Res 21(3), 199-249 (2005).

[28] V. Marian, H.K. Blumenfeld, M. Kaushanskaya, The Language Experience and Proficiency Questionnaire (LEAP-Q): assessing language profiles in bilinguals and multilinguals. J Speech Lang Hear Res 50(4), 940-967 (2007). Available at: https://bilingualism.northwestern.edu/leapq/.

[29] A. Delorme, S. Makeig, EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J Neurosci Methods 134(1), 9-21 (2004). Available at: https://sccn.ucsd.edu/eeglab/index.php.

[30] MathWorks, Inc., MATLA8B 8.0 and Statistics Toolbox 8.1, Natick, Massachusetts, United States (2020). Available at: https://www.mathworks.com/downloads/.

[31] R Core Team, R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria (2020). Available at: https://www.r-project.org/.

[32] D. Makowski, The Psycho Package: An Efficient and Publishing-Oriented Workflow for Psychological Science. Journal of Open Source Software 3(22), 470 (2018). Available at: https://github.com/neuropsychology/psycho.

[33] W. Klimesch, EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. Brain Res Brain Res Rev 29(2-3), 169-195 (1999).

[34] N.A. Herweg, E.A. Solomon, M.J. Kahana, Theta Oscillations in Human Memory. Trends Cogn Sci 24(3), 208-227 (2020).

[35] R. Erlam, Elicited imitation as a measure of L2 implicit knowledge: an empirical validation study. Applied Linguistics 27(3), 464–491 (2006).

[36] N. Spada, J.L.J. Shiu, Y. Tomita, Validating an elicited imitation task as a measure of implicit knowledge: comparisons with other validation studies. Lang Learn 65(3), 723-751 (2015).

[37] L. Pion-Tonachini, K. Kreutz-Delgado, S. Makeig, ICLabel: An automated electroencephalographic independent component classifier, dataset, and website. Neuroimage 198, 181-197 (2019).

[38] O.M. Bazanova, Alpha EEG activity depends on the individual dominant rhythm frequency. J Neurother 16(4), 270-284 (2012).