INVESTIGATING SECOND LANGUAGE LEARNING AND MUSICAL ABILITY: AN ERP STUDY

by

Kara Narzikul

Bachelor of Philosophy, University of Pittsburgh, 2015

Submitted to the Undergraduate and Graduate Faculty

of the College of Arts and Sciences in fulfillment of the

requirements for the degree of the Bachelor of Philosophy.

University of Pittsburgh

2015

UNIVERSITY OF PITTSBURGH

DIETRICH SCHOOL OF ARTS AND SCIENCES

This thesis was presented

by

Kara M. Narzikul

It was defended on

April 3, 2015

and approved by

Scott Fraundorf, University of Pittsburgh, Department of Psychology

L. Robert Slevc, University of Maryland, Department of Psychology

Alba Tuninetti, University of Pittsburgh, Department of Psychology

Thesis Advisor: Natasha Tokowicz, University of Pittsburgh, Department of Psychology

INVESTIGATING SECOND LANGUAGE LEARNING AND MUSICAL ABILITY: AN ERP STUDY

Kara M. Narzikul, BPhil

University of Pittsburgh, 2015

Individual differences such as working memory and musical ability are associated with L2 learning (e.g., Slevc & Miyake, 2006). We explored the extent to which musical ability/experience and L1-L2 similarity related to L2 learning using event-related brain potentials (ERPs). Sessions 1 and 2 included L2 grammar and vocabulary training, Sessions 2 and 3 included grammaticality assessments with ERP recording.

Goldsmith's Musical Sophistication Index (a measure of subjective musical ability on which higher scores suggest higher musical sophistication) correlated positively with grammaticality judgment test performance. ERP data revealed that scores on the Musical Ear Test (MET) for Melody (a test in which participants judged whether or not two similar melodies were the same or different) were related to a reversed N400 ERP component, but only in the first post-test. The N400 is implicated in the processing of meaning, with greater mean amplitudes suggesting greater difficulty with processing and more interference from L1 (Kutas & Federmeier, 2011). These data suggest that, initially, individuals who are more musically talented melodically have an advantage in L2 learning. Finally, there is a large early difference between individuals as a function of MET-Melody and MET-Rhythm scores. Individuals with higher MET-Melody and higher MET-Rhythm scores show a significantly more positive going waveform peaking at 100 ms, suggesting that individuals who are more musically talented

process stimuli differently than individuals who are not musically talented, perhaps indicating an attempt to convert visual stimuli into sounds.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

Thank you to Scott Fraundorf for serving on my committee.

Thank you to Alba Tuninetti for serving on my committee, and for helping me in the ERP laboratory.

Thank you to L. Robert Slevc, for serving on my committee and for providing the musical tasks that were used in this research.

Thank you to Tessa Warren for helping me throughout my proposal process.

Thank you to the University of Pittsburgh Honors College for funding this research project.

Thank you to both the graduate and undergraduate members of the PLUM Lab for their help in the ERP laboratory.

Finally, a very special thank you to my thesis advisor, Natasha Tokowicz, who has fully supported me throughout this entire process, from start to finish.

1.0 INTRODUCTION

Previous research has illustrated that the similarities between an individual's first language (L1) and second language (L2) impact L2 acquisition, and that similarities between L1 and L2 relate to neurocognitive activity during L2 processing (e.g., n, 2005; Tokowicz & MacWhinney, 2005; Tolentino & Tokowicz, 2014). More specifically, studies have used functional magnetic resonance imaging (fMRI) and event-related potential (ERP) techniques to show the impact of cross-language similarity on L2 learning. Tolentino and Tokowicz (2011) reviewed articles on this topic and found that the brain activity of L2 learners is not significantly different than the brain activity of native L2 speakers when processing cross-linguistically similar constructions. Additionally, Tolentino and Tokowicz concluded that when processing cross-linguistically unique or dissimilar constructions, nonnative learners exhibit brain activity that is different than that of native L2 speakers, and that L2 learners also show differences in the processing of dissimilar and unique features as compared to L1.

Factors at various levels are likely to influence L2 learning; namely, language-specific factors such as cross-language similarity, instructional factors, and individual differences. Cross-language similarity has been shown to influence L2 learning; however, it is only one of the many important factors to consider (MacWhinney, 2005). Instructional methods including contrast highlighting and instructional methods including both contrast highlighting and rule explanations have also been shown to facilitate L2 learning (Tolentino & Tokowicz, 2014). Finally, research has also shown that individual differences such as working memory and musical ability are associated with L2 learning (e.g., Linck, Osthus, Koeth, & Bunting, 2013; Slevc & Miyake, 2006). Most studies examine the role of factors from one or perhaps two of these levels (cross-language similarity, instructional factors, and individual differences). The present study is unique because it examined all three simultaneously, in a multi-longitudinal design. Although the primary interest in the present study was musical ability as related to cross-language similarity, this results of this research also explore the interactions between factors at multiple levels.

Language and music are both universal systems of separate, hierarchically classified elements, and past neuroimaging and event-related potential studies have shown that music and language may be processed in similar areas of the brain (Levitin & Menon, 2003; Patel, 2003; Patel, Gibson, Ratner, Besson, & Holcomb, 1998). Because this research emphasizes the importance of the shared structure and underlying neural mechanisms of music and language, it is reasonable to infer that proficiency in musical ability may be related to L2 learning.

In a study that manipulated both linguistic and musical expectations, researchers found evidence that music and the processing of syntactic structures share resources (Patel, 2003; Slevc, Rosenberg, & Patel, 2009). Slevc et al. asked participants to read garden path sentences. In a garden path sentence, structurally unexpected words often require readers to reprocess the sentence, making syntactic processing more challenging. Participants read the garden path sentences at their own pace while simultaneously listening to musical chord progressions; components of each chord were paired with each sentence segment (Slevc et al., 2009). The pairing of unexpected words with unexpected chord components resulted in enhanced garden path effects. Interestingly, this enhancement was not observed when participants encountered unexpected words, or when participants encountered unexpected chord components. This research supports the idea that language and music share underlying neural processing mechanisms.

Past research has also found links between musical ability and one specific aspect of L2 learning, phonological ability (Posedel, Emery, Souza, & Fountain, 2011). Researchers trained native English-speaking undergraduates in Spanish, ultimately testing them on pitch perception ability and Spanish pronunciation ability. The researchers used the Wing Measures of Musical Talents to gauge participants' pitch perception abilities. In this task, participants judged the differences between pairs of tones, chords, and melodies. Participants' productive phonology was evaluated based on their performance on an aural Spanish reading task in which their general pronunciation was assessed. The researchers found that participants with better pitch perception abilities exhibited more accurate Spanish pronunciation, revealing that musical ability (specifically pitch perception) is related to better L2 pronunciation.

Some previous studies that have investigated the relationship between self-reported musical ability and L2 proficiency did not have any notable findings to report (e.g., Flege, Yeni-Komshian, & Liu, 1999; Thompson, 1991). To explore this with more objective tests, Slevc and Miyake (2006) investigated the influence of musical aptitude in L2 learning by using sections from a reputable standardized test (Wing Measures of Musical Talents) to measure musical ability (Flege et al., 1999; Wing, 1968). They investigated four domains of L2 ability: receptive phonology, productive phonology, syntax, and lexical knowledge, and tested both productive and receptive phonological abilities at the word, sentence, and passage levels. Receptive phonological tests asked participants to discriminate similar-sounding L2 words and identify mispronounced words, and productive phonological tests assessed participants' general production ability (pronunciation, intelligibility, and prosody) and their ability to pronounce similar-sounding L2 words. Slevc and Miyake found that in adult L2 learners, participants' receptive and productive phonological skills are related to their musical abilities. These research findings emphasize that musical ability is associated with L2 learning, specifically in success on assessments of phonological skill.

Previous research has also compared the performance of native tone language speakers and native non-tone language speakers on assessments of learning proficiency. More specifically, Cooper and Wang (2012) compared speakers of Thai (a tonal language) with speakers of English (a non-tonal language). Participants were divided into musicians and non-musicians, based on self-reported years of musical experience, after which they received Cantonese training. Because Cantonese, like Thai, is a tonal language, Cooper and Wang examined which group(s) performed the best on tests of Cantonese tone and vocabulary. Participants also performed a musical aptitude test (The Advanced Measures of Music Audiation) to examine musical ability. Cooper and Wang found that separately, musical experience or tone language experience predicts improved non-native tone language learning ability. Interestingly, native Thai speakers who were also classified as musicians did not outperform the other groups that were solely musical or Thai-speaking, and musical ability was associated with better performance than tone language experience in tone identification. Lastly, Cooper and Wang found that scores on the tone identification and musical aptitude tasks predicted Cantonese word learning ability in the native English speakers, but not in the native Thai speakers.

Past studies on L2 learning have shown that instruction method has the ability to influence initial L2 acquisition (Schmidt, 1990). The present study was based on a previous study that investigated how the technique of L2 instruction interacts with similarities and differences between individuals' L1 and L2, while also examining the neural activity during the initial learning processes (Tolentino, 2010). In the previous study, native English speakers were taught Swedish vocabulary and grammar of three different levels of L1-L2 similarity (similar, dissimilar, and unique) and with three different learning methods. The three different learning methods included one condition with contrast and highlighting, one condition with contrast, highlighting, and added descriptions of relevant grammatical rules, and a control condition without any highlighting or rule explanations (see Tolentino & Tokowicz, 2014, for details). After training, participants completed both immediate and delayed post-tests with simultaneous recording of event-related potentials (ERPs) both to measure overall performance and to explore underlying neural activity. Results showed improvement on all groups' post-test performance; however, the control group scored the lowest on these tests (specifically for dissimilar features), whereas the other two groups scored the highest on these tests (specifically for similar and unique features). Additionally, across the different instruction groups, ERPs revealed qualitative neural differences that varied with degree of cross-language similarity (Tolentino & Tokowicz, 2010).

The present study built on the previous research described above by assessing the connection between individual differences specifically in terms of musical ability and experience and L2 learning ability. Despite the evidence for a relationship between musical ability and language proficiency, musical ability and L2 learning have remained largely unexamined. The present study sought to investigate how musical ability/experience is related to L2 learning of cross-linguistically similar, different, and unique constructions. Although the present study used the same three instructional training conditions as the previous study, due to a small number of participants in each training group we were not able to examine whether this relationship was stronger under certain instructional conditions; this question would be interesting to pursue in future research. The results of the present study explore the extent to which musical ability/experience and L1-L2 similarity relate to L2 learning in this sample, and also examine participants' concurrent neural processes through the collected event-related brain potential data.

In this study, two measures were used to gauge participants' musical ability/experience. The first is Goldsmiths Musical Sophistication Index (Gold-MSI; Müllensiefen, Gingras, Musil, Stewart, & Williamson, 2014). The Gold-MSI is a questionnaire used to measure participants' subjective musical ability. It incorporates five subscales in testing: active musical engagement, self-reported perceptual abilities, musical training, self-reported singing abilities, and sophisticated emotional engagement with music. It was administered as a paper-and-pencil task, and then scored according to the Gold-MSI guidelines. See the Appendix for the questionnaire items.

The second is the Musical Ear Test (MET; Wallentin, Nielsen, Friis-Olivarius, Vuust, & Vuust, 2010). This is a behavioral task that assesses musical abilities. The MET can classify a group into professional musicians and non-musicians, and the test can properly sort groups of professional musicians, amateur musicians, and non-musicians (Wallentin et al., 2010).

2.0 METHOD

2.1 PARTICIPANTS

Participants were 21 native English speakers recruited from the University of Pittsburgh and surrounding areas. The group consisted of 18 females and 3 males. Participants did not have any understanding of Swedish, Danish, Norwegian, German, or Dutch. Additionally, participants were not exposed to any languages, excluding English, before they turned 13. Participants ranged in age from 18 to 41 years old (M = 19 years old). The participants attended three sessions each and were compensated 10 dollars per hour for their time. Each participant received his or her total payment at the conclusion of the third session. The study included only right-handed participants with normal or normal-corrected vision.

2.2 PROCEDURE

The present study employed a 3 x 2 within subjects design with cross-language similarity (similar, dissimilar, and unique) and post-test number (post-test 1, post-test 2). The between-subjects variable manipulated in the previous study on which this study was based was learning method (control, rule, or rule and salience); this variable is not analyzed in the present study.

The ERP data that were collected focused on grammaticality (grammatical and ungrammatical), laterality (left, middle, and right), and electrode site (frontal, central, and parietal). The present study investigated the relationship of musical ability to the three factors described above.

Participants attended three separate sessions, each of which was about 1.5 to 2 hours in length. In the first session, participants completed a pre-test and received initial training in Swedish vocabulary and grammar. In the second session, participants were again trained in Swedish vocabulary and grammar. Additionally, participants completed a grammaticality judgment post-test (GJT) to measure comprehension and a sentence translation task to measure production ability. In the third session, participants completed a GJT post-test and a sentence translation task to measure the respective areas mentioned above.

More specifically, in the first session, which lasted about 1.5 hours, participants completed a filler operation-span task and then a GJT pre-test, to assess their current familiarity with Swedish (LaPointe & Engle, 1990). Finally, at the conclusion of the first session, participants also completed the Gold-MSI and the MET.

After the pre-test, participants were pseudorandomly assigned to one of the three training groups (Salience, Rule & Salience, or Control). This assignment was made considering their pre-test d-prime scores with the goal of matching the three groups. This matching was successful ($F(2,16) = 1.79$, $MSE = 0.12$, $p = .20$). Finally, participants received Swedish language and

grammar training and then completed a Language History Questionnaire (from Tokowicz, Michael, & Kroll, 2004).

About two days after the first session (between two to three days, $M = 2.10$ days) the second session occurred, which lasted about 2.5 hours. In the second session, participants completed the same vocabulary training task as in session one. Following this task, they completed an L2-L1 vocabulary translation task. After this test, participants completed the same grammar training that they completed in the first session, which again corresponded to their respective training conditions. After this, participants completed a post-test with ERP recording, after which they completed an L1-L2 sentence translation task. The final activity in the second session required participants to answer a brief questionnaire about general patterns they noticed during the GJT, handedness, medicine use, and drug use.

About two weeks after session 2 (between 14 to 17 days, M = 14.20 days), session 3 occurred. In the third session, which lasted about 1.5 hours, participants completed another post-test with ERP recording, after which they completed their last L1-L2 sentence translation task. As their final task, participants completed the Modern Language Aptitude Test "Words-In-Sentences" (MLAT-WIS) subtest as a filler task (Robinson, 1997).

2.3 MEASURES

The experiment used 372 Swedish sentences ranging from two to eight words that varied in degree of L1-L2 similarity, which were divided into training sentences and test sentences. Training included 60 sentences; the group of 60 was evenly divided into cross-language similarity type. For the post-tests, 288 sentences were split evenly, so that each post-test contained 96 sentences. In the pre-test, 24 remaining sentences were used, 12 of which were also used in all post-tests as practice sentences. The training sentences included thirty-five Swedish words, all of which were articles, nouns, verbs, adverbs, or adjectives.

In the translation task, the experiment used 24 grammatical English sentences. Post-tests 1 and 2 contained 12 sentences each. The 12 sentences were divided into sections specific to each cross-language similarity type (similar, dissimilar, and unique).

Cross-language similarity was measured using three distinct sets of Swedish training stimuli, each of which represented similar, dissimilar, or unique L1-L2 features (Tolentino & Tokowicz, 2014). Examples of each appear below. A morphosyntactic feature shared by Swedish and English is considered similar, whereas a morphosyntactic feature that occurs in both languages but is instantiated differently in each is classified as dissimilar. For example, both Swedish and English include demonstrative determiner-noun number agreement (Example 1); however, noun phrase definitiveness marking differs in Swedish in comparison to English (Example 2). In English, definitiveness is indicated by the article alone, whereas in Swedish, definitiveness is indicated on both the article and on the noun as a postfix (e.g., “pojken” = “boy-the”). More specifically, in Example 2, the sentence “The boy is eating,” is shown. Because English does not indicate definitiveness on the noun (“pojke,”), if this sentence is translated into Swedish, the sentence could become ungrammatical. Lastly, unique features are those that are present in L2 but not in L1, such as definite article-adjective gender agreement (Example 3). Swedish nouns are classified into two different genders, common or neuter. Common nouns include “en,” an indefinite article (en pojke [“a boy”]), whereas neuter nouns include the article “ett” (ett djur [“an animal”]). Unlike the English language, Swedish language rules dictate that the articles and adjectives that refer to nouns must agree with the noun in gender (e.g., A young animal” [“Ett$_{NEUT}$ ungt$_{NEUT}$ djur$_{NEUT}$”]).

The sentence examples (from Tolentino & Tokowicz, 2014) appear below:

(1) Similar (demonstrative determiner-noun number agreement): a. Den där$_{SING}$ pojken$_{SING}$ äter. [*That boy is eating*.] b. *De där$_{PL}$ pojken$_{SING}$ äter. [**Those boy is eating*.]

(2) Dissimilar (singular noun phrase definiteness): a. Pojken$_{DEF}$ äter. [*The boy is eating*.] b.*En$_{INDEF}$ pojken $_{DEF}$ äter. [**A$_{INDEF}$ boy$_{DEF}$ is eating*.]

(3) Unique (indefinite singular article-adjective gender agreement): a. En$_{COM}$ ung$_{COM}$ pojke äter. [*A young boy is eating*.] b. *En$_{COM}$ ungt$_{NEUT}$ pojke äter. [**A$_{COM}$ young$_{NEUT}$ boy is eating*.]

Using E Prime, the participants completed the GJT pre-test, judging thelye grammaticality of Swedish sentences that appeared in the center of the computer screen without receiving any feedback. They indicated their responses using two buttons.

In the vocabulary training task, participants saw an English word appear on the computer screen with its grammatical category in parentheses. The participants also heard the word presented aurally. After this, the word's Swedish translation appeared on the bottom half of the screen and was presented aurally by a native Swedish speaker. Participants were told to say the Swedish words aloud twice before pressing a button to progress to the next word. The training was designed to last 20 minutes, during which time 35 words were repeatedly shown to participants. The same vocabulary training was given in both sessions 1 and 2.

Participants were trained in Swedish grammar using one of the three assigned techniques. Each group received randomly-selected, grammatical Swedish sentences. The sentences only contained vocabulary from the vocabulary training. Using E Prime, the sentences were shown to the participants in the center of the screen and aurally. Participants needed to press a button to remove the fixation cross from the middle of the screen, causing a new sentence to be presented. They were instructed to look for grammatical patterns and to read the sentences aloud after the visual and aural presentation, before continuing to the next screen. The three training conditions, the control, rule, and rule and salience groups, were used during the grammatical training. In each condition, one third of the sentences were dissimilar, one third were similar, and one third were unique (see the description above). Differences in L2 instruction method were instantiated by three different computerized grammar training groups (Control, Salience, and Rule & Salience) each of which differed in the salience of the stimuli and the amount of additional information that was provided (Tolentino & Tokowicz, 2014). The control group viewed pairs of L2 sentences that appeared without any highlighting or explanations (e.g., “Pojken äter.” [“The boy eats/is eating.”] and “Flickan springer.” [“The girl runs/is running.”]). The salience group viewed the same sentences; however, these sentences were paired to emphasize the different instantiations of a particular morphosyntactic feature. Additionally, the sentences included bolded font and blue coloring to highlight the morphemes at the points of agreement (e.g., Definiteness marking: “**En** pojk**e** äter.” [“A boy is eating.”] vs. “Pojke**n** äter.” [“The boy is eating.”]). Finally, the rule and salience group was shown sentences with the same highlighting and coloring as the salience group, but with additional explanations of grammatical rules (e.g., “Definiteness is marked by attaching “(e)n” or “(e)t” to the end of a noun without the preceding articles “en/ett.”). The computer program was designed so that it lasted 40 minutes and the same grammatical training was used in sessions 1 and 2; however, in the first session, because of a technical issue, the first two subjects who were assigned to the Control condition received shorter training than the other control subjects (a 20 minute training session instead of a 40 minute training session). This issue was resolved for all future sessions after the second control subject completed the first session.

In the vocabulary translation task, participants completed an L2-L1 translation of words randomly chosen from the training set. The fixation cross appeared in the center of the screen, and participants needed to press a button in order for the words to appear. At the presentation of each word, participants said the English translation aloud, or said, “I don’t know.” After the verbal response was obtained, the fixation cross appeared and the process repeated.

In the GJT post-test with ERP recording, participants determined the grammaticality of sentences. With E Prime, sentences were presented on a computer screen one word at a time. The fixation cross appeared on the screen and did not disappear until the participant pressed a button. Each word was presented for 450 ms, with 350 ms pauses between the words. At the conclusion of the sentence, the participant saw a question mark appear on the screen. This question mark did not disappear until the participant responded. Depending on the response, “Correct!” or “Incorrect” appeared on the computer screen for 1000 ms after the participant’s response. Sentences were shown one word at a time so that the data did not include eye movements from reading the sentences. Before beginning the test, participants had 12 trials. After the trials, the task took about 20 minutes to complete. All GJT post-tests shared this procedure but differed in the sentences that were shown.

In the sentence translation task, participants were given an Excel spreadsheet on a computer screen with English sentences. They were asked to type the Swedish translation of each provided sentence without special characters. The same task was used in sessions 2 and 3, but with different sentences for each test.

The MET test asked participants to complete 104 trials to determine whether a pair of sounds was identical. The test incorporated melody and rhythm. It was divided into two equal sections with 52 excerpts each. In the melodic portion, participants listened to piano melodies, with three to eight components, whereas in the rhythmic portion, participants listened to woodblock rhythms, with four to eleven components. Participants marked their answers on an answer sheet. Each subtest included 26 musical phrases that were identical, and 26 musical phrases that were not, in a random order.

Electrophysiological activity was recorded during GJT post-tests at a rate of 1000 Hz. The EEG signal was amplified and participants wore electrode caps with 64 electrodes in addition to two hanging electrodes on the left and right mastoid bones. Hanging electrodes were also be placed above and below the left eye and in the left and right outer canthi. For each stimulus type, the data were calculated by using the mean amplitude, as compared to the 100 ms pre-stimulus baseline. Specifically, data were collected from 300 to 500 ms (N400/LAN) and 500 to 700 ms (LAN/P600).

Neuroscan Edit 4.3 Software was used to process the EEG data (Compumedics NeuroScan, Inc.). The electrodes were re-referenced to the average of the left and right mastoids, then low-pass filtered at 30 Hz, and finally corrected for ocular and movement artifacts. We did not include channels with large artifacts in the analyses, and two participants were excluded because of large artifacts. In order to use a participant's data, at least 8 trials (both grammatical and ungrammatical trials) needed to be artifact-free for the similar, dissimilar, and unique types. A total of 19 participants had usable ERP data. The range of the ERP epoch began at the 100 ms pre-stimulus baseline and concluded at 1000 ms post-stimulus. The mean amplitude was calculated according to this range in the 300 to 500 ms window (N400/LAN) and in the 500 to 700 ms window (LAN/P600).

3.0 RESULTS

3.1 DESCRIPTIVE STATISTICS

The three measures of musical ability/experience were not significantly intercorrelated (see Table 1).

3.2 BEHAVIORAL DATA

To account for the possibility of response bias in the behavioral data, accuracy scores from the pre-tests and post-tests were converted to d-prime scores to measure grammatical discriminability. D-prime scores range from 0 to 6, where 0 suggests no sensitivity and 6 suggests perfect sensitivity (Green & Swets, 1974). We considered correlations of d-prime scores in both post-test 1 and post-test 2 with all three measures of musical ability/experience, MET-Melody, MET-Rhythm, and Gold-MSI.

The overall Gold-MSI scores were positively correlated with the d-prime scores in both post-tests. In the data from post-test 1, there was a significant positive relationship between d-prime scores and overall Gold-MSI scores ($r = .47$, $p < .05$), and in the data from post-test 2, there was a significant positive relationship between d-prime scores and overall Gold-MSI scores ($r = .54$, $p < .05$). Additionally, there was a significant relationship between the Musical Training subscale of the Gold-MSI (see Appendix) and d-prime scores in post-test 2 (r = .46, p < .05). There was also a marginally significant relationship between the General Musical Sophistication factor, which is comprised of questions from several different subscales (see Appendix), and d-prime scores in post-test 2 (r = .49, p = .05). There were not any other significant results.[1]

3.3 ERP DATA

We examined ERP data from two time windows, corresponding to the N400/LAN component (300-500 ms) and the P600 component (500-700 ms). In each case, ERP data were analyzed from 9 electrodes, using two factors, laterality (left, midline, and right) and site (frontal, central, and parietal). The laterality factor grouped the electrodes into the following groups: left (F3, C3, P3), midline (FZ, CZ, PZ), and right (F4, C4, P4). An ANCOVA was conducted using all of the following factors: similarity, post-test, grammaticality, laterality, site, and musical test scores; continuous musical test scores were included as a covariate. Although ERP analyses were conducted separately for MET-Melody, MET-Rhythm, and MSI, we only discovered a significant relationship between MET-Melody and the ERP data, and only for the N400 time window.[2]

---

[1] Correlations run separately by similarity condition did not reveal a clear pattern of influence of cross-language similarity on the relationship between musical ability/experience and performance.
[2] There was an exception to this, which was a significant five-way interaction between post-test, cross-language similiarity, grammaticality, site, and MSI for the P600 time window. However, follow-up tests to probe this interaction did not reveal interpretable results. For example, there was only one marginally significant correlation between MSI and P600 amplitude (for post-test 1, different grammatical stimuli at parietal sites, r = .45, p = .05).

Statistical analyses revealed several significant interactions: post-test and grammaticality ($F(1,17) = 7.56$, $MSE = 42.83$, $p < .05$); post-test, grammaticality, and MET melody ($F(1,17) = 6.03$, $MSE = 42.83$, $p < .05$); post-test, grammaticality, and laterality ($F(2,34) = 4.65$, $MSE = 1.60$, $p < .05$). These three interactions were qualified by a significant four-way interaction between post-test, grammaticality, laterality, and MET melody ($F(2,34) = 5.40$, $MSE = 1.60$, $p < .01$). No other effects were significant.

To probe the four-way interaction, we split participants into relative higher and lower MET melody groups and conducted separate ANOVAs for each post-test with relative MET melody, grammaticality, and laterality as factors. The lower MET melody group (n = 8) scored between 30 and 32 and the higher group (n = 8) scored between 34 and 46. Three participants with scores in the middle range were not included in the analysis.

For post-test 1, the ANOVA revealed: a significant effect of grammaticality ($F(1,14) = 9.59$, $MSE = 3.16$, $p < .01$); a significant interaction between grammaticality and MET-Melody group (high or low) ($F(1,14) = 4.77$, $MSE = 3.16$, $p < .05$); a significant effect of laterality ($F(2, 28) = 6.83$, $MSE = 1.93$, $p < .01$); and a significant interaction between grammaticality, laterality, and MET-Melody group ($F(2,28) = 5.95$, $MSE = 0.17$, $p < .01$). For post-test 2, a significant effect of laterality was found ($F(2,28) = 9.78$, $MSE = 2.37$, $p < .01$). There were not any other significant effects.

These analyses demonstrate that in participants with lower MET-Melody scores in post-test 1, over the right, left, and middle areas, there is a more *positive going* mean-amplitude for ungrammatical stimuli (see Figure 1). In participants with higher MET-Melody scores in post-test 1, over the right hemisphere, there is a more positive going mean-amplitude for ungrammatical stimuli (see Figure 2). Interestingly, this positive going mean-amplitude effect was not significant within the similar condition alone (all *p*s $> .11$). Figure 3 shows grand average ERP waveforms as a function of grammaticality and relative MET-Melody score.

Although we did not make predictions regarding early, sensory components, it is obvious from visual inspection of the waveforms that there is a large early difference between individuals with high MET-Melody scores, and individuals with low MET-Melody scores. Specifically, individuals with high MET-Melody scores show a significantly more positive going waveform peaking at 100 ms ($F(4,68) = 3.43$, $MSE = 0.46$, $p < .05$). Additionally, individuals with high MET-Rhythm scores show the same significant effect ($F(4,68) = 3.18$, $MSE = 0.47$, $p < .05$). Finally, individuals with high Gold-MSI scores show a marginally significant relationship to the P100 component ($F(4,68) = 2.79$, $MSE = 0.48$, $p = .07$). This positive-going waveform appears again at 800 ms when the next word appears on the screen. Because of this consistent pattern, it appears that the participants with high MET-Melody and high MET-Rhythm scores have a larger sensory component whenever words appear on the screen, most evident over the parietal lobe.

4.0 DISCUSSION

Although this study did not focus on instruction method, the findings of this research differed from Tolentino's (2010) study. Specifically, Tolentino found a significant effect of similarity; similar stimuli were associated with the highest d-prime scores, dissimilar stimuli were associated with the second-highest d-prime scores, and finally, unique stimuli were associated with the lowest d-prime scores. Tolentino also found an interaction between post-test, similarity, and instruction. Although it is possible that the present study did not find such results because the posttests were restructured and instruction was not examined, it is difficult to explain these results.

In the present study, in the data from both post-tests, there was a significant relationship between d-prime scores and Gold-MSI scores. As explained above, the Gold-MSI is a subjective measure of musical ability. It measures individuals' frequency of participating in musical activities (such as listening to music, receiving formal musical training, or playing an instrument). It is divided into question subscales that assess the following: active engagement with music, perceptual abilities, musical training, singing abilities, and emotions. The Gold-MSI also includes a General Musical Sophistication measure, which contains aspects of each of the five subscales (for a total of 18 questions). In both post-tests, higher overall Gold-MSI scores were associated with higher d-prime scores, or higher accuracy on tests of grammaticality discriminability. Because formal training, engagement with, and participation in musical activities are specific actions that increase an individual’s exposure to music, the Gold-MSI measures musical interest (as opposed to musical ability, which would be measured by the MET). The correlation between Gold-MSI scores and d-prime scores occurred in both post-tests, which indicates that musical interest makes a lasting contribution to individual’s performance on GJT tests, at least within the bounds of what was measured in this study.

One possible explanation for the positive correlation between Gold-MSI and d-prime scores is that those with a greater interest in learning music may also have a greater interest in learning language. However, if that were true, we should have found a correlation between d-prime scores and scores on the engagement subscale. Research has demonstrated that music and language share resources in the processing of syntactic structures, so perhaps the participants who are more familiar with processing musical structures are also more familiar and better able to process grammatical and ungrammatical structures (Patel, 2003; Slevc, Rosenberg, & Patel, 2009). This theory also corresponds with the finding that the correlation between the Gold-MSI scores and d-prime scores was observed in both post-test 1 and post-test 2. Additionally, this theory corresponds with the finding that the Musical Training subtest was significantly related to post-test 2 scores. Perhaps general subjective musical ability (as broadly measured by the entire MSI) provides a consistent advantage throughout testing (both post-tests), whereas formal musical training and general musical sophistication provide an advantage that is more visible in later testing.

From these results, it seems as though participating in formal musical training and general musical sophistication are the more important individual differences in L2 learning. If a person spends more of his or her time participating in musical activities (specifically, musical training), according to these results, he or she will outperform another person who does not spend any time doing such activities. It is possible that the relationship between participating in musical activities and performance on grammaticality judgment tests is causal, although this study is not set up to test this idea.

With respect to the MET-Melody and MET-Rhythm, these two tests were unrelated to participants' d-prime scores; however, we found that the MET-Melody was related to participants' online processing (specifically observed in post-test 1 and via reverse N400 effects). Statistical analyses revealed a significant four-way interaction between post-test, grammaticality, laterality, and MET melody. So, interestingly, only MET-Melody was related to ERP processing and only in the first post-test. But why are only the MET-Melody scores associated with ERP processing, rather than both the MET-Melody scores and the MET-Rhythm scores? Previous research has linked language and melody, showing a strong correlation between individuals' pitch processing in both language and music (Perrachione, Fedorenko, Vinke, Gibson, & Dilley, 2013).

In participants with lower MET-Melody scores in post-test 1, over the right, left, and midline areas, there is a more positive going mean-amplitude trend for ungrammatical stimuli (see Figure 1). Tolentino (2010) also observed this reverse N400 effect in dissimilar ungrammatical stimuli. In participants with higher MET-Melody scores in post-test 1, over the right hemisphere, there is a more positive going mean-amplitude trend for ungrammatical stimuli (see Figure 2). Tolentino (2010) also observed this reverse N400 effect in dissimilar ungrammatical stimuli. Fittingly, the reverse N400 effect observed in this study is not significant within the similar stimuli alone.

Tolentino (2010) described this reverse N400 effect as interference from an individual's L1. The N400 is implicated in the processing of meaning, with greater mean amplitudes suggesting greater difficulty with processing and more interference from L1 (Shea, 2014; Tolentino, 2010). As mentioned above, in these results, a significant N400 effect was not observed within the similar stimuli alone, revealing that the similar stimuli did not elicit this interfering effect. Because the similar stimuli are morphosyntactic features shared by Swedish and English, this lack of interference is appropriate. By applying Tolentino's interference theory to these data, it suggests that initially, those who are more musically talented (with melody in particular) have an advantage in L2 learning (less interference from L1) during the second session (which includes grammar and vocabulary training) and for the dissimilar and unique stimuli (which are typically considered to be more difficult compared to the similar stimuli). However, by the second post-test (which was given in the third session, without any additional grammar or vocabulary training), this advantage disappears and is no longer significant after the two-week period. It is also interesting to note the large differences between the mean amplitudes in post-test 2. Participants with lower MET-Melody scores exhibit a greater difference between the mean amplitudes of both the grammatical and ungrammatical types over all areas of the brain. This suggests also that those with higher MET-Melody scores are also initially more consistent in their grammaticality judgments, because there is not as wide of a range.

As for the large P100 effect, past research has shown that motor training can affect a person's P100 by making it earlier or larger (Jin et al., 2010; Zwierko et al., 2014). It is possible that those who are more musically talented are processing the stimuli differently than those who are not musically talented, resulting in the P100 effect. This theory is supported by the finding that both individuals with high MET-Melody scores and individuals with high MET-Rhythm scores show a significant P100 effect. Additionally, this theory is supported by the finding that individuals with high Gold-MSI scores show a marginally significant relationship to the P100.

Because this large sensory component is in some way related to each of the three measures of musical ability, this indicates that the musically talented participants likely processed the stimuli differently than the non-musical participants. More specifically, because the participants heard audio during their training sessions, perhaps musically talented individuals attempted to convert the visual stimuli into auditory stimuli.

5.0 CONCLUSION

In regards to the musical component of this study, individual differences, such as musical ability, are related to L2 learning. Past research has linked music and language, showing that the two may be processed in similar areas of the brain (Levitin & Menon, 2003; Patel et al., 1998). This research built on a previous study and examined how musical ability relates to adult second language learning. For the behavioral data, we found that Gold-MSI scores were positively correlated with participants' d-prime scores. In this sample, the test of subjective musical ability (the Gold-MSI) was related to a participant's ability to judge whether or not a sentence was grammatical. For the ERP data, we found a four-way interaction between post-test, laterality, grammaticality, laterality, and MET scores (melody). Finally, this research revealed an interesting P100 effect, suggesting that those who are more musically talented process stimuli differently than those who are not musically talented. Together, all of these findings support: 1) the relationship between individual abilities and the likelihood of success in L2 learning and 2) the idea that shared skills, at some level, are associated with an individual's ability to learn L2.

REFERENCES

Cooper, A., & Wang, Y. (2012). The influence of linguistic and musical experience on Cantonese word learning. *The Journal of the Acoustical Society of America*, *131*(6), 4756-4769.

Ellis, N. (2006). Selective attention and transfer phenomena in L2 acquisition: Contingency, cue competition, salience, interference, overshadowing, blocking, and perceptual learning. *Applied Linguistics, 27*, 164-194.

Flege, J. E., Yeni-Komshian, G. H., & Liu, S. (1999). Age constraints on second-language acquisition. *Journal of Memory and Language, 41*, 78-104.

Green, D.M. & Swets, J.A. (1974). *Signal Detection Theory and Psychophysics.* New York: Wiley.

Jin, H., Xu, G., Zhang, J., Ye, Z., Wang, S., Zhao, L., Lin, C., & Mo, L. (2010). Athletic training in badminton players modulates the early C1 component of visual evoked potentials: A preliminary investigation. *International Journal of Psychophysiology*, *78*(3), 308-314.

LaPointe, L. B., & Engle, R. W. (1990). Simple and complex word spans as measures of working memory capacity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 1118–33.

Levitin, D. J., & Menon, V. (2003). Musical structure is processed in "language" areas of the brain: A possible role for Brodmann area 47 in temporal coherence. *NeuroImage, 20*, 2142-2152.

Linck, J. A., Osthus, P., Koeth, J. T., & Bunting, M. F. (2014). Working memory and second language comprehension and production: A meta-analysis. *Psychonomic Bulletin & Review, 21*, 861-883.

MacWhinney, B. (2005). A unified model of language acquisition. In J. Kroll & A. de Groot (Eds.). *Handbook of Bilingualism: Psycholinguistic approaches*. New York, NY: Oxford University Press.

Müllensiefen, D., Gingras, B., Musil, J., & Stewart, L., & Williamson, V. J. (2014). The Musicality of Non-Musicians: An Index for Assessing Musical Sophistication in the

General Population. *PLoS ONE, 9,* 1-23. doi: 10.1371/journal.pone.0101091

Patel, A. D. (2003). Language, music, syntax and the brain. *Nature Neuroscience*, *6*, 674-681.

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*, 717-733.

Perrachione, T. K., Fedorenko, E. G., Vinke, L., Gibson, E., & Dilley, L. C. (2013). Evidence for shared cognitive processing of pitch in music and language. *Plos ONE*, *8*(8), 1-9. doi:10.1371/journal.pone.0073372

Posedel, J., Emery, L., Souza, B., & Fountain, C. (2011). Pitch perception, working memory, and second language phonological production. *Psychology of Music, 40*, 508-517. doi: 10.1177/0305735611415145

Robinson, P. (1997). Generalizability and automaticity of second language learning under implicit, incidental, enhanced, and instructed conditions. *Studies in Second Language Acquisition*, *19*, 223-47.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, *11*, 129-158.

Shea, N. M. (2014). *Extending the N400 to the spatial domain*. (Dissertation). Retrieved from ProQuest. (1572497)

Slevc, L.R., & Miyake, A. (2006). Individual differences in second-language proficiency: Does musical ability matter? *Psychological Science, 17,* 675–681.

Slevc, L., Rosenberg J., Patel A. (2009). Making psycholinguistics musical: self-paced reading time evidence for shared processing of linguistic and musical syntax. Psychonomic Bulletin & Review, *16*, 374-381.

Thompson, I. (1991). Foreign accents revisited: The English pronunciation of Russian immigrants. *Language Learning*, *41*, 177–204.

Tokowicz, N., & MacWhinney, B. (2005). Implicit and explicit measures of sensitivity to violations in second language grammar. *Studies in Second Language Acquisition*, *27*, 173-204.

Tokowicz, N., Michael, E., & Kroll, J. F. (2004). The roles of study-abroad experience and working-memory capacity in the types of errors made during translation. *Bilingualism: Language and Cognition*, *7*, 255-272.

Tolentino, L. C. (2010). *Second language Swedish morphosyntactic instruction and cross-language similarity: An ERP investigation* (Unpublished doctoral dissertation). University of Pittsburgh, Pittsburgh, Pennsylvania.

Tolentino, L. C., & Tokowicz, N. (2011). A review of the role of cross-language similarity in L2 (morpho)syntactic processing as revealed by fMRI and ERP methods. *Studies in Second Language Acquisition, 33,* 91-125.

Tolentino, L. C., & Tokowicz, N. (2014). Cross-language similarity modulates effectiveness of second language grammar instruction. *Language Learning, 64,* 279-309.

Wallentin, M., Nielsen, A. H., Friis-Olivarius, M., Vuust, C., & Vuust, P. (2010). The Musical Ear Test, a new reliable test for measuring musical competence. *Learning and Individual Differences, 20,* 188–196.

Wing, H. (1968). Tests of musical ability and appreciation: An investigation into the measurement, distribution, and development of musical capacity (2nd ed.). London: Cambridge University Press.

Zwierko, T, Lubinski, W., Lesiakowski, P., Steciuk, H., Piasecki, L., & Krzepota, J. (2014). Does athletic training in volleyball modulate the components of visual evoked potentials? A preliminary investigation. *Journal of Sports Medicine, 32* (16), 1519-1528.

APPENDIX A

GOLD-MSI SUBSCALES

Note: Participants were asked to circle one of the following responses to each question: 1 = Completely Disagree; 2 = Strongly Disagree; 3 = Disagree; 4 = Neither Agree nor Disagree; 5 = Agree; 6 = Strongly Agree; 7 = Completely Agree.

Active Engagement (9 items)
I don’t spend much of my disposable income on music
I enjoy writing about music for example on blogs and forums
I have attended ___ the following number of live music events as an audience member in the past twelve months
I keep track of new music that I come across e.g. new artists or recordings
I listen attentively to music for ___ per day
I often read or search the internet for things related to music
I spend a lot of my free time doing music related activities
I’m intrigued by musical styles I’m not familiar with and want to find out more
Music is kind of an addiction for me - I couldn’t live without it
Perceptual Abilities (9 items)
I am able to judge whether someone is a good singer or not
I can compare and discuss differences between two performances or versions of the same piece of music
I can tell when people sing or play out of time with the beat
I can tell when people sing or play out of tune
I find it difficult to spot mistakes in a performance of a song even if I know the tune
I have trouble recognizing a familiar song when played in a different way or by a different performer
I usually know when I’m hearing a song for the first time
When I hear music I can usually identify its genre
When I sing I have no idea whether I’m in tune or not

Musical Training (7 items)
I engaged in regular, daily practice of a musical instrument (including voice) for ___ years
I can play the following number of musical instruments (including voice)
I have never been complimented for my talents as a musical performer
At the peak of my interest, I practised ___ hours on my primary instrument (including voice)
I have had formal training in music theory for ___ years
During my life time I have had formal training on a musical instrument (including voice) for ___ years
I would not consider myself a musician
Singing Abilities (7 items)
After hearing a new song two or three times, I can usually sing it by myself
I am able to hit the right notes when I sing along with a recording
I am not able to sing in harmony when somebody is singing a familiar tune
I can sing or play music from memory
I don't like singing in public because I'm afraid that I would sing wrong notes
I only need to hear a new tune once and I can sing it back hours later
If somebody starts singing a song I don't know, I can usually join in
Emotions (6 items)
I am able to identify what is special about a given musical piece
I am able to talk about the emotions that a piece of music evokes in me
I often pick certain music to motivate or excite me
I sometimes choose music that can trigger shivers down my spine
Music can evoke my memories of past people and places
Pieces of music rarely evoke emotions for me
General Musical Sophistication (18 items)
I would not consider myself a musician
I engaged in regular daily practice of a musical instrument (including voice) for ___ years
I have never been complimented for my talents as a musical performer
I can sing or play music from memory
At the peak of my interest I practiced ___ hours on my primary instrument (including voice)
I am able to hit the right notes when I sing along with a recording
I spend a lot of my free time doing music related activities
Music is kind of an addiction for me - I couldn't live without it
I don't like singing in public because I'm afraid that I would sing wrong notes
When I sing I have no idea whether I'm in tune or not
After hearing a new song two or three times, I can usually sing it by myself
I can play the following number of musical instruments (including voice)
I only need to hear a new tune once and I can sing it back hours later
I often read or search the internet for things related to music
I am able to identify what is special about a given musical piece
I am not able to sing in harmony when somebody is singing a familiar tune
I enjoy writing about music for example on blogs and forums

I can compare and discuss differences between two performances or versions of the same piece of music

APPENDIX B

TABLES AND FIGURES

Table 1: Correlations between Measures of Musical Experience/Ability

Measure	1	2	3
1. MET-Melody	-		
2. MET-Rhythm	.37	-	
3. Gold-MSI	.06	.36	-

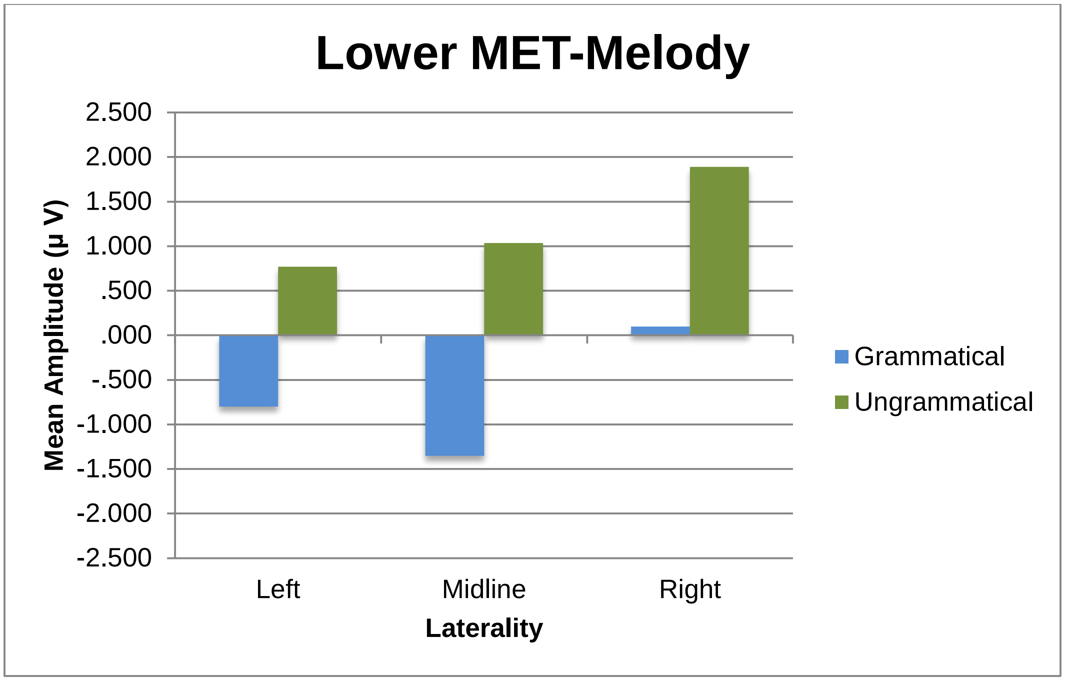

Figure 1: This figure illustrates part of the 4-way interaction in the N400 (300 to 500 ms) time window.

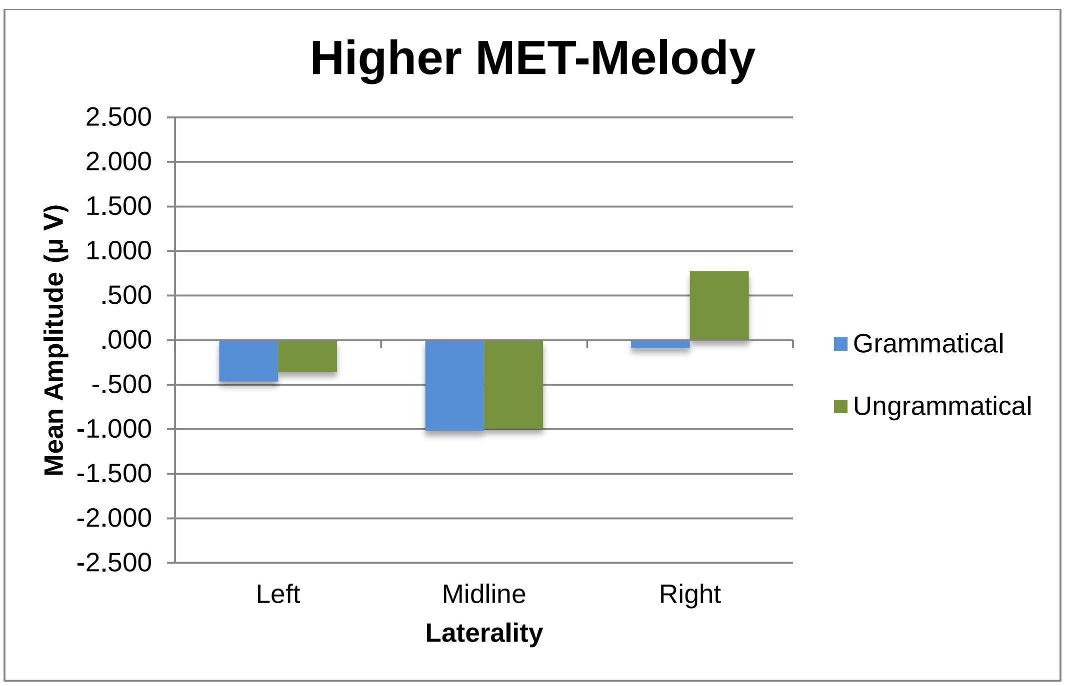

Figure 2: This figure illustrates part of the 4-way interaction in the N400 (300 to 500 ms) time window.

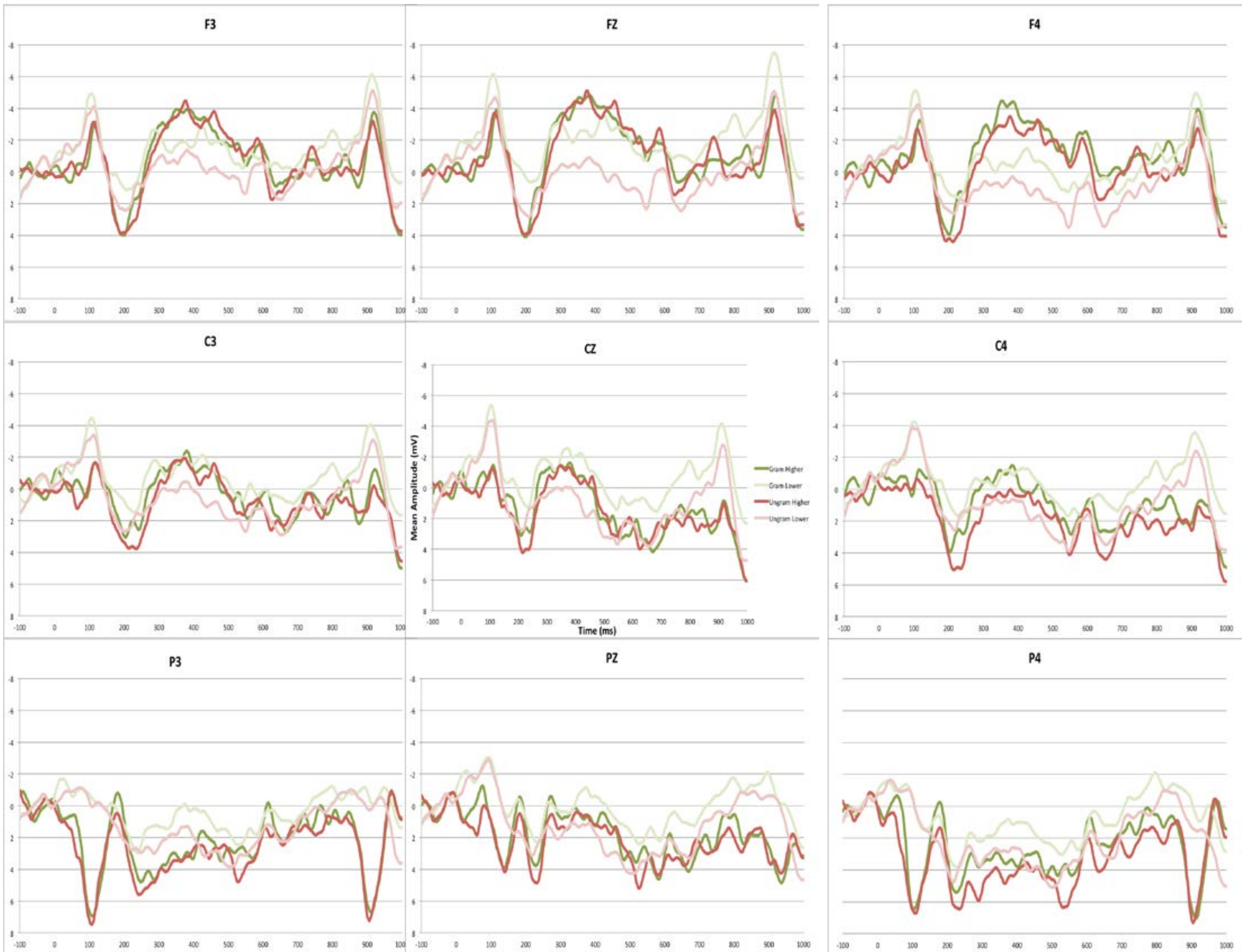

Figure 3: This figure illustrates the ERP waveforms for the higher MET-Melody and lower MET-Melody groups by grammaticality.