The Effect of Music Experience on Auditory Sequential Learning: An ERP Study

Samantha N. Emerson (semerson2@student.gsu.edu)
Department of Psychology, P.O. Box 5010, Atlanta, GA 30302 USA

Jerome Daltrozzo (jdaltrozzo@gsu.edu)
Department of Psychology, P.O. Box 5010, Atlanta, GA 30302 USA

Christopher M. Conway (cconway@gsu.edu)
Department of Psychology, P.O. Box 5010, Atlanta, GA 30302 USA

Abstract

The existence of an advantage in sequential learning for musicians over nonmusicans is highly debated. The current study used an auditory sequential learning task to investigate the neurophysiological correlates of sequential learning in adults with either high or low music aptitudes. While behavioral results alone revealed no difference between the reaction times of the two groups, event-related potential data showed that higher music aptitude was associated with decreased amplitudes of the P300 and Contingent Negative Variation effect between two conditions with different transitional probabilities relative to a target stimulus. These data suggest that increased music training and skill leads to more efficient processing of (i.e., reduced attentional demands for) auditory sequential patterns.

Keywords: Sequential learning; music; event-related potentials (ERP); P300; CNV

Introduction

Sequential learning (SL) is the ability to either implicitly or explicitly extract statistical probabilities from series of discrete elements and to form expectations based on that probabilistic information (Conway & Christiansen, 2001; Conway & Pisoni, 2008). This skill is particularly important to the development of language and has been implicated in the acquisition of word boundaries (Saffran, Aslin, & Newport, 1996), syntax (Ullman, 2004), and word order (Conway, Bauernschmidt, Huang, & Pisoni, 2010). The role of experience in shaping SL mechanisms is still relatively underspecified. Conway, Pisioni, Anaya, Karpicke, and Henning (2011) found that an early period of sound deprivation (i.e., in children with cochlear implants) led to deficits in SL abilities. On the other hand, it is possible that increased experience or skill with sound—for example music—might lead to an *advantage* in SL.

Sequential Learning in Musicians

The existence of an advantage in SL for musicians over nonmusicians is highly debated. For example, Rohrmeier, Rebuschat, and Cross (2011) found that musicians did not show an advantage over nonmusicians in their familiarity with musical sequences produced from an artificial grammar. Similarly, Bigand's (2003) review of the literature strengthens this view by demonstrating similarities in performance between musicians and nonmusicians in the processing of melodic and harmonic structures, in the processing of large-scale structures, and in implicit learning for musical structures. Bigand argues that nonmusicians' every day exposure to music makes them "expert listeners" and therefore are as competent as musicians with respect to the implicit understanding of the complex structures—or grammars—that underlie music.

This behavioral (as well as some neural) evidence appears to suggest that musical expertise does not improve SL abilities in the auditory domain. However, recent neural studies support the view of an advantage in SL with increased musical expertise. For example, two Magnetoencephalographic (MEG) studies (Herholz, Boh, & Pantev, 2011; Paraskevopoulos, Kuchenbuch, Herholz, & Panteva, 2012) examined exposure to deviant sequences of tones embedded within more standard sequences. Both studies found that musicians and nonmusicians responded similarly to the deviant sequences. However, Herhloz et al. (2011) found that only musicians exhibited an increased mismatch negativity (MMN; Näätänen & Alho, 1995) within 10 minutes of exposure, and Paraskevopoulos et al. (2012) found a significantly larger amplitude of the P50 in comparison to nonmusicians. These results demonstrate an effect of musical expertise on pre-attentive auditory abilities and short-term auditory learning of statistical regularities.

An electrophysiological study examined musicians and nonmusicians' learning of statistical regularities in a sung "language" in which each syllable was associated with a particular note (Francois & Schön, 2011). Behavioral analysis revealed that both musicians and nonmusicians were able to segment the syllables and notes based on the musical structure. Electroencephalographic (EEG) analysis, however, revealed that, compared to nonmusicians, musicians exhibited a larger N1 component and a larger negativity in the 750 to 850 ms latency band in response to untrained linguistic segments. For untrained musical segments, compared with nonmusicians, musicians exhibited larger N1 and P2 components which were larger over the left hemisphere than the right, a negativity in the 350 to 500 ms latency band which was largest over the central and left frontal regions (i.e., an N400-like effect), and a negativity in the 700 to 800 ms latency band which was larger over the right frontal region. These results appear to indicate that musicians have more robust representations of both musical and linguistic structures.

The Current Study

The present study examined the relationship between SL of auditory stimuli and music aptitude through the use of an Event-Related Potential (ERP) paradigm based on a visual SL task created by Jost, Conway, Purdy, and Hendricks (2011). The task involved the presentation of a series of complex tones wherein target tones could be predicted with varying levels of probability by the preceding tone. ERPs were compared across two levels of music aptitude with three different types of predictors reflecting high, low, and zero probability of being followed by the target.

Methods

Participants

A total of 13 participants (1 male, 1 left-handed, M_{age} = 21.30, SD_{age} = 2.63) from Georgia State University participated in the study to receive class credit. Participants were divided into two groups according to their music aptitudes as determined by the Advanced Measures of Music Audiation (AMMA; Gordon, 1989). Seven participants were placed into the low music aptitude group with a total AMMA score that placed them lower than the 57th percentile (M_{AMMA} = 46.57, SD_{AMMA} = 11.15) and six participants were placed into the high music aptitude group with a total AMMA score that placed them higher than the 58th percentile (M_{AMMA} = 73.17, SD_{AMMA} = 8.86). The two groups did not differ in mean age (*F*(1, 12) = 0.22, *p* = .650, η_p^2 = .019; Table 1).

Participants were given a music questionnaire to assess different features of their musical backgrounds, and a series of one-way ANOVAs were performed on the data (Table 1) to determine whether the participants differed with respect to these features. The analyses revealed that the two groups differed significantly in both the maximum number of years they had played an instrument or sung (*F*(1, 11) = 7.08, *p* = .024, η_p^2 = .414) and in the maximum number of years of training they had received (*F*(1, 12) = 7.37, *p* = .020, η_p^2 = .401). These differences demonstrate that the high music aptitude group had a richer musical background than the low music aptitude group in terms of the number of years both playing and training on a musical instrument or voice.

Participants were also given a series of standardized measures including the Peabody Picture Vocabulary (PPVT-IV; Dunn & Dunn, 2007), the Grammaticality Judgment and Sentence Completion subtests of the Comprehensive Assessment of Spoken Language (CASL; Carrow-Woolfolk, 1999), and the Perceptual Reasoning subtests of the Wechsler Abbreviated Scale of Intelligence (WASI-II; Wechsler, 2011; Table 1). A series of one-way ANOVA's demonstrated that the two groups did not differ in their receptive vocabularies (PPVT-IV; *F*(1, 12) = 0.60, *p* = .812, η_p^2 = .005), abilities to identify and correct grammatical errors (GramJudg; *F*(1, 12) = 2.51, *p* = .141, η_p^2 = .186); abilities to accurately complete sentences (SentComp; *F*(1, 12) = 0.37, *p* = .557, η_p^2 = .032), or nonverbal perceptual reasoning (PRI; *F*(1, 12) = 1.81, *p* = .206, η_p^2 = .141). The results demonstrate that participants were well matched for these features across the two groups.

Table 1: Means (SD) for music questionnaire and assessments for music aptitude groups[1].

	Low	High
Age	20.79 (1.97)	21.68 (3.22)
Playing	3.25 (3.77)	11.50 (7.09)
Training	1.75 (2.06)	8.83 (6.52)
PPVT-IV	105.57 (9.00)	106.83 (10.26)
GramJudg	90.71 (18.33)	104.17 (10.44)
SentComp	91.86 (30.26)	100.50 (18.61)
PRI	95.43 (8.06)	109.00 (25.44)

Stimuli

A total of nine complex tones were generated for the study using Praat version 5.3.42. The seven non-target tones had a fundamental frequency ranging from 200 Hz to 1400 Hz in 200 Hz increments with two harmonics: one which was twice the frequency of the fundamental and one that was three times. The two target tones had fundamental frequencies of 2400 Hz and 2600 Hz and were constructed in a similar manner as non-targets. All tones were root mean squared normalized and presented at an intensity of 72 dBA, as measured by a digital sound level meter (Extech Instruments—Model 407732) in the fast (125 ms) and low intensity (35 to 100 dB) acquisition modes. Stimuli had a duration of 50 ms and were presented with a stimulus onset asynchrony of 1 second.

Experimental Paradigm

Participants listened to a series of complex tones and were asked to press a button on a button box whenever they heard one of two high pitched "target tones". For each participant, one of the seven non-target tones was pseudo-randomly chosen as a "standard" tone. Each trial began with the standard tone repeating a random number of times. Next, one of three "predictor" tones was played. One tone was pseudo-randomly selected for each participant to be a "high predictor" tone, another as a "low predictor" tone, and the standard tone was repeated as the "zero predictor" tone. In

---

[1] *Playing* and *Training* were measured as the maximum number of years reported for playing or receiving training on any one instrument or voice; *PPVT-IV* was measured as the standardized PPVT-IV score; *GramJudg* and *SentComp* were the standardized scores of the Grammaticality Judgment and Sentence Completion subtests of the CASL; *PRI* was the Perceptual Reasoning Index based on the Block Design and Matrix Reasoning subtests of the WASI-II.

the high predictor condition, the predictor tone was followed by a target in 90% of the trials and by the standard tone in 10% of the trials. In the low predictor condition, the predictor was followed by a target in 20% of the trials and by the standard in 80% of the trials. In the zero predictor condition, the predictor tone (which was indistinguishable from the standard tone) was followed by one of the targets in 50% of the trials and by a non-target tone (again randomly chosen for each participant) in the other 50%. Because each trial started with a random length of standard tones, the occurrence of the target could not be predicted by any occurrence of the standard tone. Each trial concluded with a second series of standards of a random length.

For each predictor condition there were 50 trials for a total of 150 trials divided amongst five blocks of 30 trials. Trials were randomly ordered across the three predictor conditions (high, low, and zero) in a continuous fashion such that the participant was unable to distinguish one trial from another. A break lasting a minimum of 30 seconds was given between each block. Stimuli were presented on a Dell Optiplex 755 computer running E-Prime version 2.0.8.90.

Recording Technique

ERP recordings were taken from 256 scalp sites using an Electrical Geodesic Inc. (EGI) sensor net (Figure 1) and were processed using Net Station Version 4.3.1. Electrode impedances were kept below 50 kΩ. Recordings were made with a 0.1 to 30 Hz bandpass filter and digitized at 250 Hz. The continuous EEG was segmented into epochs -200 ms to +1500 ms with respect to the predictor onset. ERPs were baseline-corrected at 200 ms prestimulus and averaged referenced. Separate ERPs were computed for each participant, predictor type, and electrode. All experimental sessions were conducted in a 132 square foot double-walled, sound-deadened acoustic chamber.

Data Analysis

Statistical calculations were performed on averaged traces from each individual on the mean amplitudes (from baseline) within time-windows of interest estimated through preliminary analyses using latency windows of 50 ms in the 0 to 1500 ms range (Schirmer & Kotz, 2003). To test the cortical distribution of the effects, nine regions of interest (ROIs) were selected as levels of a topographic within-subjects factor based on three levels of laterality and antero-posteriority (Figure 1). Mixed-measures ANOVAs were performed on EEG means with the following factors: Predictor Type (high, low, zero), Music Aptitude (high, low), Antero-posteriority (anterior, central, posterior), and Laterality (left, middle, right). All reported *p*-values were adjusted with the Greenhouse–Geisser correction for non-sphericity, when appropriate. Scheffé's tests were used for post hoc comparisons. The reported partial eta squared is a measure of effect size for ANOVAs (Cohen, 1988; Olejnik & Algina, 2003). The statistical analyses were conducted with Cleave (January 30, 2005 Version). Cleave automatically performed all Scheffé's post hoc tests (Šídák corrected for multiple comparisons) on the significant main effects and on all significant interactions.

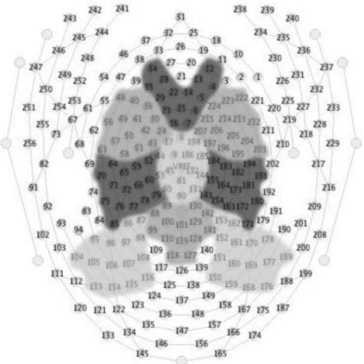

Figure 1: 256 sensors EEG net with the nine regions of interest highlighted.

Results

Behavioral Results

A mixed-measures ANOVA was conducted on the reaction times with Predictor Type (high, low, zero) as within-subjects factor and Music Aptitude (high, low) as a between-subjects factor. A main effect of Predictor Type was found ($F(2, 80) = 7.14$, $p = .005$, $\eta_p^2 = .417$). Contrasts revealed that participants were faster to respond to the targets in the high predictor condition ($M = 394.67$ ms, $SD = 76.38$) than in the zero predictor condition ($M = 455.31$ ms, $SD = 67.74$, $p = .024$), indicating that learning had occurred.

These behavioral results indicate that with exposure, participants were more able to extract the statistical structure embedded within the sequences such that they were quicker to respond to a target after it followed a high predictor tone than when it did not. Behavioral data did not reveal a difference in reaction times between participants with high and low music aptitude.

ERP Results

Sequential Learning Figure 2 displays the grand average ERPs across all participants for each of the three predictor conditions in the nine ROIs. Visual inspection suggests that there is an increased positivity for both the high and low predictor conditions in the 300 to 400 ms latency range, especially in the centro-posterior sites, followed by an increased negativity in the 500 to 1000 ms latency range. The negativity effect appears to be larger in the anterior and medial regions. A series of mixed-measure ANOVAs were performed on latency ranges of interest identified by the preliminary analyses to verify these observations.

First, a main effect of Predictor Type was found for three latency ranges: 150 to 250 ms ($F(2, 20) = 10.88$, $p = .001$, $\eta_p^2 = .497$), 300 to 450 ms ($F(2, 20) = 10.79$, $p = .001$, $\eta_p^2 = .495$), and 650 to 1250 ms ($F(2, 20) = 13.39$, $p = .001$,

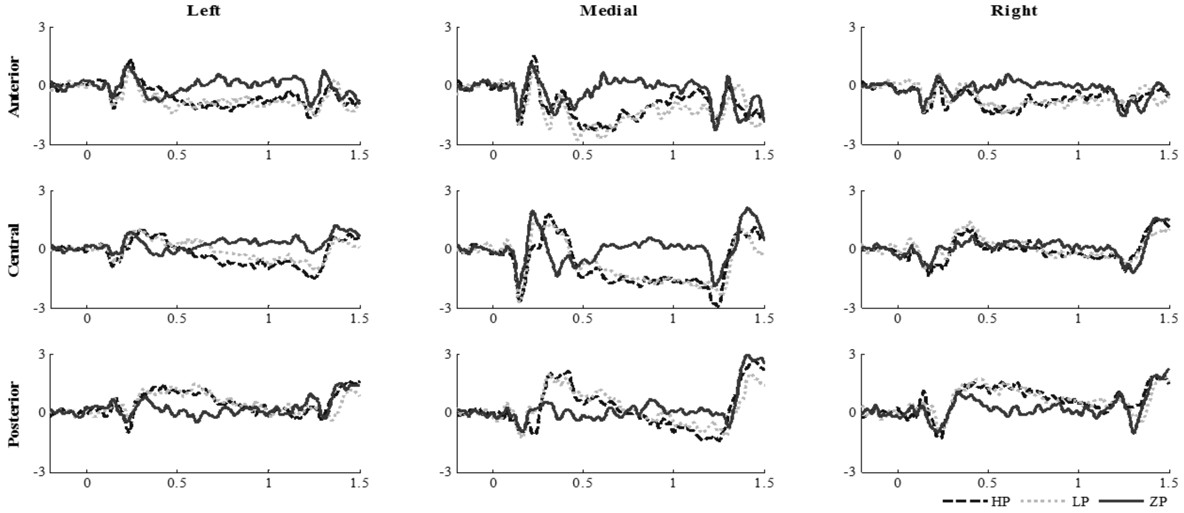

Figure 2: Grand average ERPs for all nine ROIs in response to the predictors in the high (*HP*), low (*LP*), and zero (*ZP*) predictor conditions (positivity upward in microVolts; time in seconds).

$\eta_p^2 = .549$). Post hoc tests for all four time-windows did not reach significance (Šídák $\alpha = .0170$).

An interaction between Predictor Type and Antero-posteriority was significant in the 550 to 850 ms time-window ($F(4, 44) = 6.24$, $p = .012$, $\eta_p^2 = .362$). Post hoc analyses did not reach significance (Šídák $\alpha = .0014$); however, the ERP effect was marginally significantly more negative in the anterior than posterior region for both the high (anterior: $M = -1.32$, $SD = 2.13$; posterior: $M = 0.72$, $SD = 1.67$, $p = .039$) and low predictor conditions (anterior: $M = -1.26$, $SD = 1.73$; posterior: $M = 0.80$, $SD = 1.43$, $p = .037$).

An interaction between Predictor Type and Laterality also reached significance in the 600 to 800 ms time-window ($F(4, 44) = 4.76$, $p = .008$, $\eta_p^2 = .302$). Post hoc tests did not reach significance (Šídák $\alpha = .0014$); however, several contrasts approached significance: In the medial region, the potential for the high ($M = -1.07$ µV, $SD = 2.34$) and low ($M = -0.93$ µV, $SD = 2.04$) predictor conditions were more negative than the zero predictor condition ($M = 0.15$ µV, $SD = 1.17$; $p = .033$, $p = .058$, respectively), and the potential for the high predictor condition was more negative in the medial region ($M = -1.07$, $SD = 2.34$) than in the right region ($M = 0.06$, $SD = 1.91$; $p = .046$).

Altogether the SL results primarily suggest the existence of a late (550-850 ms) fronto-central negativity in both the high and low predictor conditions as compared to the zero predictor. The neural responses to the high and low predictor conditions did not differ from one another, suggesting that both patterns had been encoded similarly despite the fact that the low predictor tone was not as consistent of a predictor as the high predictor tone.

Music Aptitude Mixed-measures ANOVAs revealed a three-way interaction between Music Aptitude, Predictor Type, and Laterality in the 300 to 400 ms time-window ($F(4, 44) = 3.49$, $p = .002$, $\eta_p^2 = .241$) and the 650 to 900 ms time-window ($F(4, 44) = 5.74$, $p = .003$, $\eta_p^2 = .343$).

For the 300 to 400 ms time-window (Figure 3), post hoc tests did not reach significance (Šídák $\alpha = .0003$); however, several contrasts approached significance: For the low predictor condition in the right region, participants in the low music aptitude group ($M = 1.26$ µV, $SD = 2.33$) showed a greater positivity than participants in the high music aptitude group ($M = -0.07$ µV, $SD = 1.15$, $p = .037$). In the medial region, participants in the low music aptitude group showed a greater positivity for the high ($M = 0.95$ µV, $SD = 3.41$) and low ($M = 0.56$ µV, $SD = 3.58$) predictor conditions than the zero predictor condition ($M = -0.90$ µV, $SD = 1.34$; $p = .004$, $p = .022$, respectively). For the low music aptitude participants, the low predictor condition was more positive in the right region ($M = 1.26$, $SD = 2.33$) than in the left region ($M = 0.02$, $SD = 2.19$, $p = .049$).

For the 650 to 900 ms time-window (Figure 3), post hoc tests revealed two contrasts that reached significance and several others that approached it (Šídák $\alpha = .0003$): In the high predictor condition, participants in the low music aptitude group showed a negativity that was marginally significantly greater than participants in the high music aptitude group in the left ($M = -0.89$ µV, $SD = 1.87$; $M = 0.23$ µV, $SD = 0.81$; $p = .043$) and medial ($M = -1.82$ µV, $SD = 2.29$; $M = -0.21$ µV, $SD = 1.01$; $p = .005$) regions. For the left region, participants in the low music aptitude group produced a negativity that was marginally significantly greater in the high predictor condition ($M = -0.89$ µV, $SD = 1.87$) than in the zero predictor condition ($M = 0.23$ µV, $SD = 0.92$, $p = .041$). For the medial region, participants in the low music aptitude group produced a negativity that was significantly greater in the high predictor condition ($M = -1.82$ µV, $SD = 2.29$) than in the zero predictor condition ($M = 0.30$ µV, $SD = 1.04$, $p < .001$) and marginally significantly greater in the low predictor condition ($M = -1.39$ µV, $SD = 2.22$) than in the zero predictor condition (M

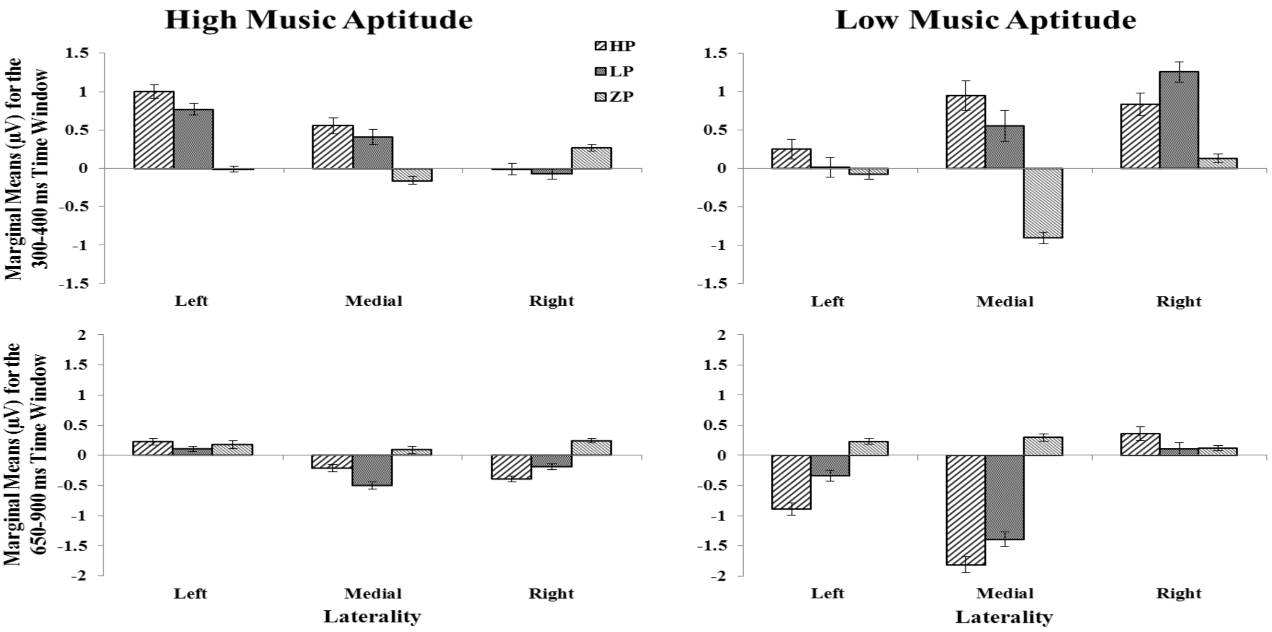

Figure 3: Marginal means (µV) for the 300-400 ms (top) and 650-900 ms (bottom) time-windows for the high (*HP*), low (*LP*), and zero (*ZP*) predictor conditions for participants in the high (left) and low (right) music aptitude groups.

= 0.30 µV, *SD* = 1.04, *p* = .003). In the low predictor condition, the participants in the low music aptitude group produced a negativity that was almost significantly greater in the medial region (*M* = -1.39, *SD* = 2.22) than in the right (*M* = 0.11, *SD* = 1.85, *p* = .008). In the high predictor condition, the participants in the low music aptitude group produced a negativity that was significantly greater in the medial (*M* = -1.82, *SD* = 2.29) than right (*M* =0.36, *SD* = 2.03, *p* < .001) region

The results for Music Aptitude analyses implicate two ERP components. In the 300 to 400 ms time-window, marginally significant larger positivities were found for the high and low predictor conditions in comparison to the zero predictor. This positivity effect was more prominent in the low music aptitude group than the high. In the 650 to 900 ms time-window, marginally significantly larger negativities were found in the low music aptitude group for the high and low predictor conditions in comparison to the zero predictor condition. In particular, the negativity effect between high and zero predictor conditions reached significance in the medial region and was greater in the medial region than in the right region.

Discussion

Analysis of reaction times revealed that participants with both high and low music aptitudes were able to extract the probabilistic relationship between the predictors and the target tones, demonstrating that both groups exhibited SL. It is important to note that if we had access to behavioral data only, we would have concluded—like previous studies (e.g., Bigand, 2003; Rohrmeier, Rebuschat, & Cross, 2011)—that music aptitude did not influence SL. However, in line with recent neurophysiological studies (Francois & Schön, 2011; Herholz et al., 2011; Paraskevopoulos et al., 2012), our ERP results revealed differences between participants with high and low music aptitude. In particular, the amplitudes of two ERP components differentiated the two groups.

The first component was a centro-parietal positive deflection occurring 300-400 ms post-predictor onset that was larger for the high and low predictor conditions as compared to the zero predictor condition. The ERP component showing this effect has the typical topography and latency of a P300 (Polich, 2007) and is likely to index the same cognitive mechanisms as the late positive component found with a similar SL paradigm but in the visual modality (Jost et al., 2011). The amplitude of this component has been suggested to index the amount of attentional resources involved in processing a stimulus (Polich, 2007; Polich & Bondurant, 1997).

The second ERP component that differentiated the music aptitude groups was a fronto-central negative deflection that occurred 650-900 ms post-predictor onset for the high and low predictor conditions as compared to the zero predictor condition. Unlike the P300 effect, this negative ERP effect was not found by Jost et al. (2011) and likely reflects a modulation of the Contingent Negative Variation (CNV; Walter, Cooper, Aldridge, McCallum, & Winter, 1964) which is thought to reflect anticipation of a stimulus. Similar to the P300, the amplitude and the latency of the CNV may be modulated by the attentional resources devoted to a particular task or stimulus (Tecce, 1972).

In the case of these two ERP components, marginally significantly greater amplitudes—for the right and medial regions in the case of the P300 and for the left and medial in the CNV—were seen for the participants with low music aptitudes compared to those with high aptitudes. This pattern of results suggests that the level of attention given to the predictor tones by the low music aptitude group was higher than that of the high aptitude group. This difference in the levels of attention between groups, despite equivalent levels of learning as measured by reaction times, may indicate a difference in cognitive effort. That is to say, the low music aptitude group—being less expert in processing sound sequences than the high music aptitude group—may have required additional cognitive resources to reach the same level of behavioral performance as the high music aptitude group, as reflected by increased ERP amplitudes.

The current paradigm cannot fully dismiss the possibility that some of the observed behavioral and ERP effects between the zero predictor condition with the high and low predictor conditions are due to an (early preattentive) orienting mechanism rather than to SL. However, the cognitive mechanisms typically associated with the CNV are not preattentive orienting mechanisms. Therefore, it is likely that the CNV effects presented here which differentiate the high from the low music aptitude groups are due to SL rather than orienting mechanisms.

In conclusion, the results of this study suggest that high music aptitude is associated with more efficient processing of sequential auditory stimuli. Because the measure of music aptitude was strongly related to the number of years of playing and receiving training in a musical instrument, it is likely that increased music experience improves the brain's efficiency at processing auditory sequential patterns.

Acknowledgments

We thank GSU's Language and Literacy Initiative and the NIH (R01DC012037) for their financial support and S. Singh and M. Freggens for their help with this project.

References

Bigand, E. (2003). More about the musical exertise of musically untrained listeners. *Annals of the New York Academy of Sciences, 999*, 304-312.

Carrow-Woolfolk, E. & American Guidance Service (1999). *Comprehensive assessment of spoken language (CASL).* Circle Pines, MN: American Guidance Service.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum Associates.

Conway, C. M., Bauernschmidt, A., Huang, S. S., & Pisoni, D. B. (2010). Implicit statistical learning in language processing: Word predictability is key. *Cognition, 114*, 356-371.

Conway, C. M. & Christiansen, M. H. (2001). Sequential learning in non-human primates. *Trends in Cognitive Sciences, 5*(12), 539-546.

Conway, C. M. & Pisoni, D. B. (2008). Neurocognitive basis of implicit learning of sequential structure and its relation to language processing. *Annals of the New York Academy of Sciences, 1145*, 113-131.

Dunn, L. M., Dunn, L. M., & Pearson Assessments (2007). *Peabody Picture Vocabulary Test-Fourth Edition*. Circle Pines, MN: American Guidance Service (AGS), Inc.

Francois, C., & Schön, D. (2011). Musical expertise boosts implicit learning of both musical and linguistic structures. *Cerebral Cortex, 21*(10), 2357-2365.

Gordon, E. E. (1989). *Advanced measures of music audiation*. Chicago, IL: G.I.A. Publications.

Herholz, S. C., Boh, B., & Pantev, C. (2011). Musical training modulates encoding of higher-order regularities in the auditory cortex. *European Journal of Neuroscience, 34*(3), 524-529.

Jost, E., Conway, C. M., Purdy, J. D., and Hendricks, M. A. (2011). Neurophysiological correlates of visual statistical learning in adults and children. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the 35th Annual Conference of the Cognitive Science Society* (pp. 2526-2531). Austin, TX: Cognitive Science Society.

Näätänen, R., & Alho, K. (1995). Mismatch negativity: A unique measure of sensory processing in audition. *International Journal of Neuroscience, 80*(1-4), 317-337.

Olejnik, S., & Algina J. (2003). Generalized eta and omega squared statistics: Measures of effect size for some common research designs. *Psychological Methods, 8*(4), 434–447.

Paraskevopoulos, E., Kuchenbuch, A., Herholz, S. C., & Pantev, C. (2012). Statistical learning effects in musicians and non-musicians: An MEG study. *Neuropsychologia, 50*(2), 341-349.

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology, 118*(10), 2128-2148.

Polich, J. & Bondurant, T. (1997). P300 sequence effects, probability, and interstimulus interval. *Physiological Behavior, 61*, 843-849.

Rohrmeier, M., Rebuschat, P., & Cross, I. (2011). Incidental and online learning of melodic structure. *Consciousness and Cognition: An International Journal, 20*(2), 214-222.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274*, 1926-1928.

Schirmer, A., & Kotz, S. A. (2003). ERP evidence for a sex-specific Stroop effect in emotional speech. *Journal of Cognitive Neuroscience 15*(8). 1135–1148.

Tecce, J. J. (1972). Contingent negative variation (CNV) and psychological processes in man. *Psychological Bulletin, 77*(2), 73-108.

Ullman, M. T. (2004). Contribution of memory circuits to language: The declarative/procedural model. *Cognition, 92*, 231-270.

Walter, W. G., Cooper, R., Aldridge, V. J., McCallum, W. C., & Winter, A. L., (1964). Contingent Negative Variation: An electric sign of sensory-motor association and expectancy in the human brain. *Nature 203*, 380–384.

Wechsler, D., & Psychological Corporation. (2011). *WASI - II: Wechsler abbreviated scale of intelligence—Second Edition*. San Antonio, Tex: Psychological Corporation.