 doi:10.20944/preprints201906.0122.v1

Attention Modulates Electrophysiological Responses to Simultaneous Music and Language Syntax Processing

Daniel J Lee[1,2], **Harim Jung**[1,3], **Psyche Loui**[1,4]

[1]Wesleyan University

[2]Columbia University

[3]University of Pennsylvania

[4]Northeastern University

Abstract

Music and language are hypothesized to share neural resources, particularly at the level of syntax processing. Recent reports suggest that attention modulates this sharing of neural resources, but the time-course of the effects of attention, and the degree to which attention operates similarly on music and language, are yet unclear. In this EEG study we manipulate the syntactic structure of simultaneously presented musical chord progressions and garden-path sentences in a modified rapid serial visual presentation paradigm, while varying top-down attentional demands to the two modalities. The Early Right Anterior Negativity (ERAN) was observed in response to both attended and unattended musical syntax violations. In contrast, an N400 was only observed in response to attended linguistic syntax violations, and a P3 only in response to attended musical syntax violations. Results show that top-down allocation of attention indeed affects the processing of syntax in both music and language, with different neural resources acting upon the two modalities particularly at later stages of cognitive processing. However, the processing of musical syntax at an earlier stage of the perceptual-cognitive pathway, as indexed by the ERAN, is partially automatic, and is strongly indicative of separate neural resources for music and language.

Keywords: music; language; syntax; attention; comprehension; electroencephalography; event-related potentials

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 June 2019 doi:10.20944/preprints201906.0122.v1

1. Introduction

Music and language are both fundamental to human experience. The two domains, while apparently different, rely on several notable similarities: both exhibit syntactic structure, and both rely on sensory, cognitive, and vocal-motor apparatus of the central nervous system. The nature of this relationship between syntactic structure in language and music, and their underlying neural substrates, is a topic of intense interest to the cognitive and brain sciences community.

The Shared Syntactic Integration Resources Hypothesis (SSIRH, Patel, 2003) is an influential theoretical account of similarities and differences between cognitive processing for music and language. The SSIRH posits that neural resources for music and language are shared at the level of the syntax; in other words, processing of music and language should interact at the syntactic level, but not at other levels such as semantics or acoustic or phonemic structure.

Support for the SSIRH comes from a variety of behavioral and neural studies. Several studies have presented music and language simultaneously with and without syntactic violations, to test for effects of separate and simultaneous syntax violations on behavioral and neural measures (Fedorenko, Patel, Casasanto, Winawer, & Gibson, 2009; Jentschke, Koelsch, & Friederici, 2005; Koelsch, Gunter, Wittfoth, & Sammler, 2005; Roncaglia-Denissen, Bouwer, & Honing, 2018; Slevc, Rosenberg, & Patel, 2009). Strong support for the SSIRH comes from a self-paced reading paradigm (Slevc et al., 2009), where sentence segments were presented concurrently with musical chord progressions. Subsets of the trials contained syntactic violations in language (garden path sentences) and syntactic violations in music (out-of-key chords). Reaction time results showed a superadditive effect, a statistical interaction between domain (music vs. language) and violation (syntactically congruent vs. incongruent). In contrast, a control experiment which manipulated the timbre of the music and the semantics of the language did not elicit the same superadditive effect.

Although these results seem to offer convincing support for SSIRH, Perruchet and Poulin-Charronnat (2013) showed that under semantic garden path manipulations (as opposed to syntactic garden path manipulations), violations of semantics can also yield the same pattern, suggesting that increased task demands, rather than syntax processing per se, could lead to these statistical interactions.

The idea that task demands can influence the pattern of interaction between music and language processing has since received more support. More recent work has argued that the sharing of processing resources of syntax for language and music might have a shared reliance on domain-general resources, such as attention, cognitive control, and decision-making strategies, especially when simultaneously processing music and language in a dual-task situation (Roncaglia-Denissen et al., 2018; Slevc & Okada, 2015).

At a neural level, the Early Right Anterior Negativity (ERAN) and the Early Left Anterior Negativity (ELAN) are reliably elicited event-related potential (ERP) markers of syntax processing in music and language respectively (Hahne & Friederici, 1999; Koelsch, Gunter, Friederici, & Schroger, 2000; Koelsch et al., 2005; Koelsch, Schmidt, & Kansok, 2002; Sammler et al., 2013). The ERAN is a frontally-generated negative waveform around 200 ms after the onset of musical syntax violations, whereas the ELAN is an analogous frontally-generated negativity after violations in linguistic syntax. Musical syntax processing has been localized in magnetoencephalographic (MEG) studies to the inferior frontal gyrus (Cheung, Meyer, Friederici, & Koelsch, 2018; Maess, Koelsch, Gunter, & Friederici, 2001), and additional results from ERAN of lesioned patients (Sammler et al., 2009; Sammler, Koelsch, & Friederici, 2010; Sun et al., 2018) have also provided causal evidence for the reliance of musical syntax on classic language-related areas such as the inferior frontal gyrus. The ERAN is also posited as an index of predictive processes in the brain, especially in the case of music, due to its reliance on the formation and subsequent violation of predictions that are learned from exposure to musical sound sequences (Koelsch, Vuust, & Friston, 2018). Importantly, the ERAN is also sensitive to top-down task demands, such as attentional resources devoted to the musical stimuli in contrast to a concurrent, non-musical task (Loui, Grent-'t-Jong, Torpey, & Woldorff, 2005). Impaired ERAN is observed in children with Specific Language Impairment, and in adults with lesions in the left inferior frontal gyrus (Broca's area) (Jentschke, Koelsch, Sallat, & Friederici, 2008; Sammler, Koelsch, & Friederici, 2011), which provides additional support for the SSIRH.

While ERAN and ELAN are markers of syntax processing, semantic processing is indicated by the N400, a negative-going centroparietal waveform beginning around 400 ms after a semantic anomaly (Kutas & Hillyard, 1980, 1984). In addition to being sensitive to semantic content of words, the N400 effect reflects the semantic associations between words and the expectancy for more generally, showing a larger waveform as an incoming word is unexpected or semantically incongruous with the previous context. In response to ambiguities in linguistic syntax that violate the ongoing context, the P600 is another effect that has also been observed, especially during garden path sentences when a newly presented word or words require a reinterpretation of the preceding context (Kuperberg, 2007). The P600 is similar in topography and latency to the P3, a complex of positive-going event-related potentials that are elicited from 300 ms and onwards following an unexpected and task-relevant event. The P3 is separable into two components: P3a, a fronto-central ERP, largest around FCz that reflects novelty processing, and P3b, a later parietally-centered ERP largest around Pz that is more sensitive to motivation and task demands (Polich, 2007). Patients with frontal lobe lesions show altered habituation of the P3, suggesting that the amplitude of the P3 is subject to frontally-mediated processes (Knight, Grabowecky, & Scabini, 1995). The P3a and P3b are both elicited in syntactically incongruous events in music, and are sensitive to different levels and genres of expertise (Koelsch et al., 2002; Przysinda, Zeng, Maves, Arkin, & Loui, 2017).

Although the electrophysiological markers of syntax processing in music and language are relatively well established, the strong predictions offered by SSIRH and their subsequent predictions on cognitive control have yet to be tested in simultaneous behavioral and EEG paradigms. Specifically, we do not yet know how the previously observed superadditive behavioral effect of musical and linguistic syntax on reaction time (Slevc et al., 2009) plays out in neurophysiological indices. A detailed characterization and comparison of ERPs, as elicited by musical and linguistic syntactically congruent and incongruent processes, will shed light on the cognitive processes underlying music and language, such as by disambiguating between early sensory expectations and further cognitive processing.

Furthermore, the question of whether neural processes are dedicated to musical and linguistic syntax, or whether syntactic processes for music and language both tap into domain-general resources for cognitive control, can be directly addressed by experimentally manipulating top-down attentional processes to music and language. By comparing electrophysiological activity during experimental conditions in which stimuli are the same, but attentional demands are differently directed towards either music or language, we aim to compare musical and linguistic syntax processing under different conditions of attentional control. Additionally, by simultaneously measuring behavioral and electrophysiological indices, we can obtain a temporally precise understanding of musical and linguistic syntax processing under different attentional conditions.

In this study we compare ERP processes elicited by violations in musical and linguistic syntax, while attention was directed separately towards language or music. We adapted Slevc et al's (2009) paradigm of simultaneous musical and linguistic syntax processing during a rapid serial presentation paradigm while

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 June 2019 doi:10.20944/preprints201906.0122.v1

the task was language comprehension, and extended this study by adding a musical analog of the language comprehension task. Thus, across two experiments we were able to compare behavioral results during task-manipulated attention to language and music, while independently manipulating syntax in each domain at a finer timescale in order to test for effects in event-related potentials that are known markers of syntax processing and attention.

2. Materials and Methods

2.1 Subjects

Thirty-five undergraduate students from Wesleyan University participated in return for course credit. All participants reported normal hearing. Informed consent was obtained from all subjects as approved by the Ethics Board of Psychology at Wesleyan University. Sixteen students (11 males and 5 females, mean age = 19.63, SD = 2.03) were assigned to the Attend-language group: 15/16 participants in this group reported English as their first language, and 9/16 participants reported prior music training (total mean of training in years = 2.23, SD = 3.42). Nineteen students (8 males and 11 females, mean age = 19.40, SD = 2.03) were assigned to the Attend-music group. Background survey and baseline tests of one participant in this group was missing as the result of a technical error. Of the remaining 18 participants, 12/18 reported English as their first language, and 11/18 participants reported having prior music training (total mean of training in years = 3.11, SD = 4.01).

The years of musical training was not different between the two groups ($X^2 = 0.0215$, $p = .88$). Table 1 shows the demographics of the participants in both conditions.

Variable	Attend-language (N = 16)	Attend-music (N = 19)
Age, year	19.625 (2.029)	19.389 (2.033)
Range, year	18-24	18-26
Male, n	11/16	8/19
Music Training, year	2.233 (3.422)	3.105 (4.012)
Trained, n	9/16	11/18
Shipley	16.125 (3.008)	16.631 (2.477)
MBEA	23.375 (3.828)	25.11 (2.685)
Pitch Perception, Hz	12.469 (11.895)	11.087 (7.174)
Normal Hearing, %	100%	100%
English as First Language, n	15/16	12/18

Table 1. Demographics of the participants. Data are shown as mean (SD), range, or proportion. SD: Standard Deviation. n: Count in proportion.

2.2 Stimuli

The stimuli were adapted from Slevc et al. (2009) (Slevc et al., 2009). 48 English sentences were simultaneously presented with 48 musical chord progressions. Each segment of a sentence was paired with a chord that followed the rules of Western tonal harmony in the key of C major, played in a grand piano timbre. Linguistic syntax expectancy was manipulated through syntactic garden-path sentences, whereas musical syntax expectancy was manipulated through chords that were either in-key or out-of-key at the highlighted critical region segment (Fig. 1). The blocks were presented at the regular inter-onset interval of 1200ms. At the end of each sentence, a yes/no comprehension question was presented on the screen: In the Attend-Language condition, this question was on the content of the sentence to direct participants' attention to language (e.g. "Did the attorney think that the defendant was guilty?"). For the Attend-Music condition, the question at the end of the trial asked about the content of the music (e.g. "Did the music sound good?") to direct participants' attention to the music. Participants' task, in both Attend-Language and Attend-Music groups, was always to respond to the question at the end of each trial by choosing "Yes" or "No" on the screen.

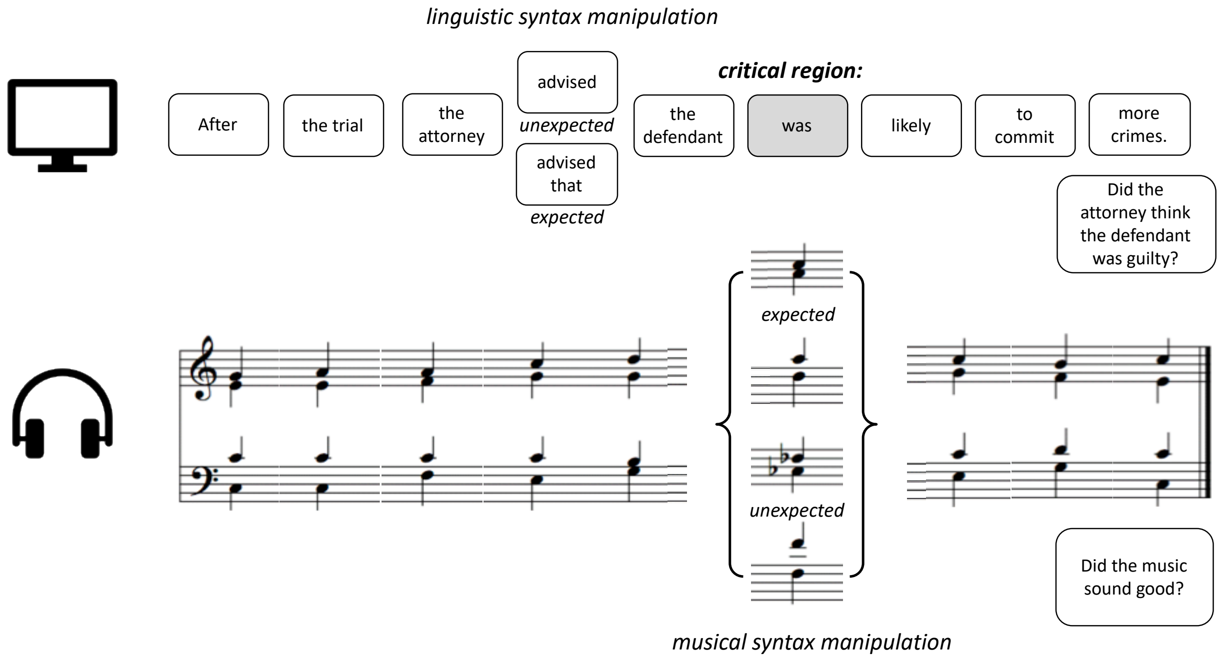

Figure 1. Example trials for Attend-language and Attend-music conditions.

2.3 Procedure

Participants first gave informed consent and filled out a background survey on their musical training, as well as Shipley, MBEA, and pitch perception baseline tests before the main experiment. Before the experiment began, the participants were told to pay attention to every trial, and to answer a yes-or-no comprehension question about the language (Attend-language condition) or about the music (Attend-music condition) at the end of each trial. They were given a short practice run of 5 trials of the experiment in order to familiarize themselves with the task before EEG recording began. EEG was recorded using PyCorder software from a 64-channel BrainVision actiCHamp setup with electrodes corresponding to the international 10-20 EEG system. Impedance was kept below 10 kOhms. The recording was continuous with a raw sampling rate of 1000 Hz. EEG recording took place in a sound attenuated, electrically shielded chamber.

2.4 Behavioral Data Analysis

Behavioral data from Max/MSP were imported to Excel to compute the accuracy of each participant. Accuracy was evaluated against 50% chance-level in one-sample two-tailed t-tests in SPSS. For experiment 2, two subjects' behavioral data were lost due to technical error.

 doi:10.20944/preprints201906.0122.v1

2.5 EEG Preprocessing

BrainVision Analyzer software (Brain Product Gmbh) 2.1 was used to preprocess raw data. EEG data were first re-referenced to TP9 and TP10 mastoid electrodes, and filtered with high-pass cutoff of 0.5 Hz, low-pass cutoff of 30 Hz, roll-off of 24 dB/oct, and a notch filter of 60 Hz. Ocular correction ICA was applied to remove eye artifacts for each subject. Raw data inspection was done semi-automatically by first setting maximal allowed voltage step as 200 µV/ms, maximal difference of values over a 200ms interval as 400 µV, and maximal absolute amplitude as 400 µV. Then, manual data inspection was performed to remove segments with noise due to physical movements.

2.6 Event-Related Potential Analysis

The preprocessed data were segmented into four conditions: music congruent, music incongruent, language congruent, and language incongruent. Each segment was 1200ms long, spanning from a 200ms baseline before the onset of the stimulus to 1000ms after stimulus onset. The segments were averaged across trials, baseline-corrected, and grand-averaged across the subjects. To identify effects specific to syntax violations in each modality, a difference wave was created for each violation condition by subtracting ERPs for congruent conditions from ERPs for incongruent conditions, resulting in a Music-specific difference wave (Music violation minus no violation) and a Language-specific difference wave (Language violation minus no violation). From these difference waves we isolated ERP amplitudes at two recording sites at two time windows of interest: E(L/R)AN from site FCz at 180-280 ms, and the N4 and P3 at site Pz at 500-600 ms. The mean amplitude of each ERP was exported for each participant from BrainVision Analyzer into SPSS for analysis.

Because both groups of participants experienced both types of syntactic violations (music and language), but each group of participants attended to only one modality (music or language), we used a mixed-effects analysis of variance (ANOVA) with the within-subjects factor of Violation (two levels: music and language) and the between-subjects factor of Attention (two levels: attend-music and attend-language). This was separately tested for the two time windows: 1) the early ERAN/ELAN time window of 180-280 ms (“Early”), and 2) the later N4/P3 time window of 500-600 ms (“Late”).

3. Results

3.1 Behavioral Results

Participants performed well above the 50% chance level on language comprehension questions during the Attend-Language condition ($M = 0.8457$, $SD = 0.0703$, two-tailed t-test against chance level of 50% correct: $t(15) = 19.661$, $p < 0.001$), and on music comprehension questions during the Attend-Music condition ($M = 0.6631$, $SD = 0.1253$, two-tailed t-test against chance level of 50% correct: $t(16) = 5.371$, $p < 0.001$). This confirms that participants successfully attended to both language and music stimuli. Participants performed better on the Attend-Language than on the Attend-Music questions ($t(31) = 5.12$, $p < .001$).

3.2 Event-Related Potentials

Effects of Violations were highly significant for both ERAN and P3 and N4, showing that syntactic violations in music and language elicited significantly different ERPs at both time windows. Interestingly, the between-subjects factor of Attention was significant only in the late time window: the effects of violations interacted with attention only for P3, but not for ERAN. Figure 2 shows each ERP and scalp

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 June 2019 doi:10.20944/preprints201906.0122.v1

topos of difference waves. Figure 3 shows the specific effects for each ERP; statistics are shown in Table 2.

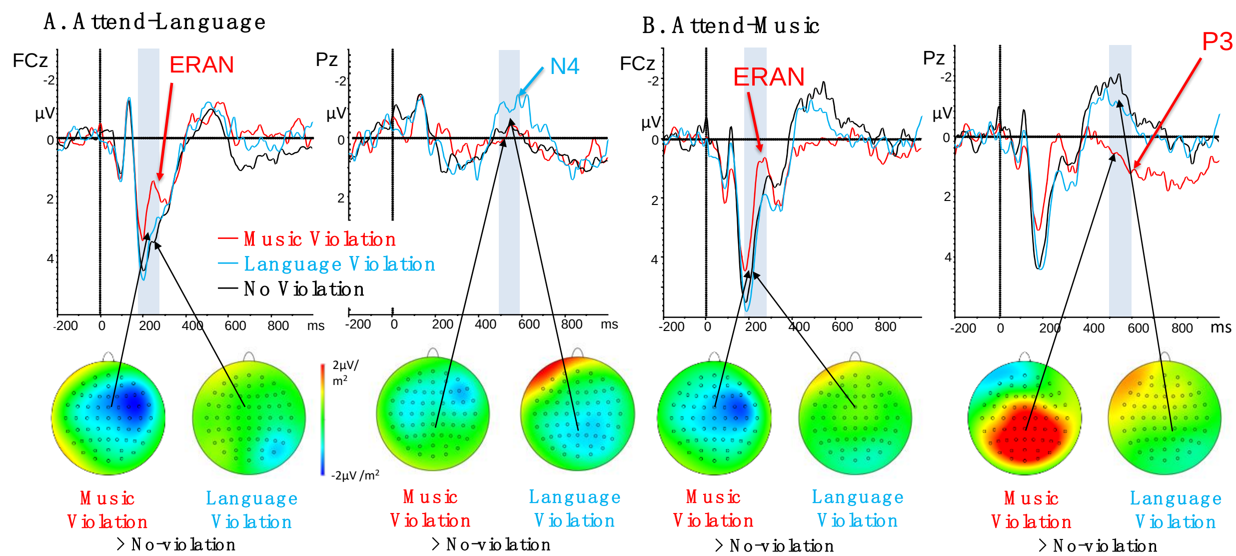

***Figure 2**. Overlays of ERPs from each condition with topographic maps of the difference wave between violation and no-violation conditions. Music syntax violation condition is shown in red and linguistic syntax violation condition is shown in blue. Black represents a condition when neither stimulus was violated. Topos plots show difference waves between music violation and no-violation, or between language violation and no-violation. **A**. When attending to language. **B**. When attending to music.*

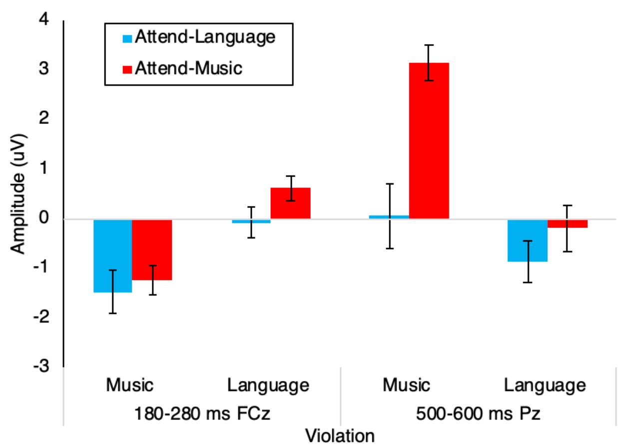

***Figure 3.** ERP effects of violation (amplitude of difference waves) across different conditions.*

Tests of Within-Subjects Contrasts

Source	Time	SS	df	MS	F	p
Violation	Early	46.403	1	46.403	33.198	< .001
	Late	79.02	1	79.02	31.317	< .001
Violation * Attend	Early	0.894	1	0.894	0.64	0.43
	Late	25.108	1	25.108	9.951	0.003
Error(Violation)	Early	46.126	33	1.398		

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 June 2019 doi:10.20944/preprints201906.0122.v1

Source	Measure	SS	df	MS	F	p
	Late	83.268	33	2.523		

Tests of Between-Subjects Effects

Source	Measure	SS	df	MS	F	p
Intercept	Early	20.769	1	20.769	7.597	0.009
	Late	19.816	1	19.816	3.164	0.084
Attend	Early	3.776	1	3.776	1.381	0.248
	Late	61.146	1	61.146	9.763	0.004
Error	Early	90.218	33	2.734		
	Late	206.684	33	6.263		

Table 2. *ERP statistics.*

A mixed-model ANOVA with a between-subjects factor of attention (two levels: attend-music vs. attend-language) and a within-subjects factor of modality of syntax violation (two levels: music and language) showed significant effects at each time-point.

180-280 ms: For the early time window at FCz, the within-subjects effect of Violation was significant ($F(1,33) = 33.198$, $p < 0.005$), whereas the between-subjects effect of Attention was not significant ($F(1,33) = 1.381$, $p = 0.248$). There was no significant interaction between the factors. This component responded significantly only to music syntax violations (Figures 2); in contrast no significant difference was observed between linguistic syntax violation and the no-violation condition (Figure 3).

500-600 ms. For the late time window at Pz, the within-subjects effect of Violation was significant ($F(1,33) = 31.317$, $p < 0.001$), and the between-subjects effect of Attention was significant ($F(1,33) = 9.763$, $p = 0.004$). Here, an Attention by Violation interaction was also significant ($F(1,33) = 9.951$, $p = 0.003$). This interaction is visible in the ERP traces and topo plots in Figure 2 as well as in the amplitude results plotted in Figure 3: in the Attend-Language condition, only language violations elicited a negative waveform resembling an N400, whereas music violations were no different from the no-violation condition. The N400 shows a latency of 400-800 ms and a centro-parietal topography (Figure 2A), consistent with classic reports of the N400 effect (Kutas and Hillyard, 1984). In contrast, during the Attend-Music condition, only music violations elicited a large positive P3 waveform, whereas language violations showed no difference from the no-violation condition. The P3 shows a latency of 400-1000 ms and a centro-parietal topography (Figure 2B), consistent with the P3b subcomponent of the P3 complex (Polich, 2007). This attention-dependent double dissociation between P3 and N400 is observable in Figures 2 and 3. In other words, the P3 was only observed for music violations when attention was directed to music, and the N400 was only observed for language violations when attention was directed to language.

4. Discussion

By separately manipulating linguistic syntax, musical syntax, and attention via task demands during simultaneous music and language processing, we were able to disentangle the effects of top-down attention on bottom-up processing of syntax and syntactic violations. Three main findings come from the current results: 1) For both music and language, syntactic violation processing activates a cascade of neural events, indexed by early and late ERP components as seen using time-sensitive methods. This replicates prior work (Koelsch et al, 2000, Hahne and Friederici, 1999, and many others). 2) Early components are less sensitive to attentional manipulation than late components, also replicating prior

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 June 2019 doi:10.20944/preprints201906.0122.v1

work (Donchin et al., 1984; Hillyard, Hink, Schwent, & Picton, 1973). 3) Attention affects musical and linguistic syntax processing differently at late time windows. This finding is novel as it extends previous work that identify early and late components in music and language syntax processing, by showing that the late components are most affected by attention. Taken together, results show that top-down manipulations of attention differently affect the bottom-up processing of music and language, with effects of attention becoming more prominent throughout the temporal cascade of neural events that is engaged during music and language processing.

In some respects, the present results add to a modern revision of the classic debate on early- vs. late-selection theories of attention. While early-selection theories (Broadbent, 1958) posited that attention functions as a perceptual filter to select for task-relevant features in the stimulus stream, late-selection theories have provided evidence for relatively intact feature processing until semantic processing (Deutsch & Deutsch, 1963; Gray & Wedderburn, 1960) or until feature integration (Treisman & Gelade, 1980). Due to their fine temporal resolution, ERP studies provide an ideal window into this debate, allowing researchers to quantify the temporal cascade of neural events that subserve perceptual-cognitive events such as pitch and phoneme perception, and syntax and semantics processing. ERP results from dual-task paradigms such as dichotic listening have shown that attention modulates a broad array of neural processes from early sensory events (Woldorff et al., 1993; Woldorff & Hillyard, 1991) to late cognitive events (Falkenstein, Hohnsbein, Hoormann, & Blanke, 1991; Näätänen, Gaillard, & Mäntysalo, 1978). Here we observe the ERAN in response to musical syntax violations regardless of whether attention was directed to language or to music. The ERAN was elicited for music violations even when in the attend-language condition; furthermore its amplitude was not significantly larger during the attend-music condition. This result differs from previous work showing that the ERAN is larger during attended than during unattended conditions (Loui et al., 2005). The difference likely stems from the fact that while in the previous study the visual task and the musical task were temporally uncorrelated, in the present study the language stimuli (sentence segments) and musical stimuli (chords) were simultaneously presented, with each language-music pair appearing in a time-locked fashion. Thus, when in the attend-language condition, the onset of musical chords became predictably coupled with the onset of task-relevant stimuli (sentence segments), even though the musical chords themselves were not task-relevant. This predictable coupling of task-irrelevant musical onsets with task-relevant linguistic stimulus onsets meant that it became more advantageous for subjects to allocate some bottom-up attentional resources to the music, or to allocate attentional resources to all incoming sensory stimuli at precisely those moments in time when stimuli were expected (Large & Jones, 1999), as one modality could help predict the other. The fact that the ERAN was observed even when only slightly attended provides some support for a partially automatic processing of musical syntax, as posited in previous work (Loui et al., 2005). When musical syntax violations were not task-relevant but were temporally correlated with task-relevant stimuli, they elicited intact early anterior negativity but no late differences from no-violation conditions. This early-intact and late-attenuated pattern of ERP results is also consistent with the relative attenuation model of attention, which posits that unselected stimulus features are processed with decreasing intensity (Treisman, 1960).

We operationalize ERAN as music-elicited negativity, and ELAN as language-elicited negativity. Here we observe the right-sided ERAN in response to musical violations, but no ELAN in response to language violations. This may be because the garden path stimuli in the language violations are too subtle, as previous studies that elicited ELAN used tense violations rather than garden path sentences (Hahne & Friederici, 1999). The introduction of the linguistic garden path requires that participants re-parse the syntactic tree structure during the critical region of the trial; this effort to re-parse the tree likely elicited

 doi:10.20944/preprints201906.0122.v1

the N4 at the later time window, but lacks the more perceptual aspect of the violation that likely elicited the ELAN in prior studies (Hahne et al, 1999).

It is remarkable that linguistic syntax violations only elicited a significant N400 effect, and no significant effects over any other time windows, even when language was attended to in Experiment 1. In contrast, musical syntax violations elicited the ERAN as well as the P3 in the attended condition, with the ERAN being observed even when musical syntax was unattended. Note that the P3 effect in this experiment is similar in topography and latency to the P600, which has been observed for semantic processing during garden path sentences. It could also be the Central-Parietal Positivity (CPP), which reflects accumulating evidence for perceptual decisions (O'Connell, Dockree, & Kelly, 2012), which can resemble the P3 (van Vugt, Beulen, & Taatgen, 2019). During the attend-music condition, linguistic syntax violations elicited no significant ERP components compared to no-violation conditions. This suggests a strong effect of attention on language processing. It is also worth noting that we saw a clear N400 and not a P600 or a P3 in response to garden path sentences in language. The relationship between experimental conditions and N400 vs. P600 or P3 is an ongoing debate in neurolinguistics: Kuperberg (2007) posits that the N400 reflects semantic memory-based mechanisms whereas the P600 reflects prolonged processing of the combinatorial mechanism involved in resolving ambiguities (Kuperberg, 2007). Others argue that whether an N400 or a P600 is observed may in fact depend on the same latent component structure; in other words. The presence and absence of N400 and P600 may reflect two sides of the same cognitive continuum, rather than two different processes per se (Brouwer & Crocker, 2017; Brouwer, Crocker, Venhuizen, & Hoeks, 2017; Brouwer & Hoeks, 2013). If the N400 and P600 are indeed two sides of the same coin, then this could mean that language and music processing are also more related than the different effects would otherwise suggest.

Limitations. One caveat is that, similar to the original paradigm which we borrow from in this study (Slevc et al., 2009), music was always presented auditorily, whereas language was always presented visually. Thus, the differences we observe between musical and linguistic syntax violation processing could also be due to differences in the modality of presentation. In future studies it may be possible to reverse the modality of presentation, such as by visually presenting musical notation or images of hand positions on a piano (Bianco et al., 2016) with spoken sentence segments. Although doing so would require a more musically trained subject pool who can read musical notation or understand the images of hand positions, prior ERP studies suggest that visually presented musical-syntactic anomalies would still elicit ERP effects of musical syntax violation, albeit with different topography and latency (Bianco et al., 2016). Furthermore, although participants performed above chance on both attend-language and attend-music comprehension questions, they did perform better on the attend-language task; this imposes a behavioral confound that may affect these results. It is yet unclear how syntax and semantics could be independently manipulated in music. In fact, changing musical syntax most likely affects the meaning participants derive from the music; however, specifically composed pieces with target words in mind might be a way to get at a musical semantics task without overtly manipulating syntax (Koelsch et al., 2004). Nevertheless, by separately manipulating music and language during their simultaneous processing, and crossing these manipulations experimentally with top-down manipulations of attention via task demands, we observe a progressive influence of attention on the temporal cascade of neural events for the processing of music and language.

 doi:10.20944/preprints201906.0122.v1

5. References

Bianco, R., Novembre, G., Keller, P. E., Scharf, F., Friederici, A. D., Villringer, A., & Sammler, D. (2016). Syntax in Action Has Priority over Movement Selection in Piano Playing: An ERP Study. *J Cogn Neurosci, 28*(1), 41-54.

Brouwer, H., & Crocker, M. W. (2017). On the Proper Treatment of the N400 and P600 in Language Comprehension. *Front Psychol, 8*, 1327.

Brouwer, H., Crocker, M. W., Venhuizen, N. J., & Hoeks, J. C. J. (2017). A Neurocomputational Model of the N400 and the P600 in Language Processing. *Cogn Sci, 41 Suppl 6*, 1318-1352.

Brouwer, H., & Hoeks, J. C. (2013). A time and place for language comprehension: mapping the N400 and the P600 to a minimal cortical network. *Front Hum Neurosci, 7*, 758.

Cheung, V. K. M., Meyer, L., Friederici, A. D., & Koelsch, S. (2018). The right inferior frontal gyrus processes nested non-local dependencies in music. *Sci Rep, 8*(1), 3822.

Deutsch, J. A., & Deutsch, D. (1963). Attention: Some theoretical considerations. *Psychological review, 70*(1), 80.

Donchin, E., Heffley, E., Hillyard, S. A., Loveless, N., Maltzman, I., Ohman, A., . . . Siddle, D. (1984). Cognition and event-related potentials. II. The orienting reflex and P300. *Ann N Y Acad Sci, 425*, 39-57.

Falkenstein, M., Hohnsbein, J., Hoormann, J., & Blanke, L. (1991). Effects of crossmodal divided attention on late ERP components. II. Error processing in choice reaction tasks. *Electroencephalography and Clinical Neurophysiology, 78*(6), 447-455.

Fedorenko, E., Patel, A., Casasanto, D., Winawer, J., & Gibson, E. (2009). Structural integration in language and music: Evidence for a shared system. *Memory & Cognition, 37*(1), 1-9.

Gray, J., & Wedderburn, A. (1960). Shorter articles and notes grouping strategies with simultaneous stimuli. *Quarterly Journal of Experimental Psychology, 12*(3), 180-184.

Hahne, A., & Friederici, A. D. (1999). Electrophysiological evidence for two steps in syntactic analysis. Early automatic and late controlled processes. *J Cogn Neurosci, 11*(2), 194-205.

Hillyard, S. A., Hink, R. F., Schwent, V. L., & Picton, T. W. (1973). Electrical signs of selective attention in the human brain. *Science, 182*(4108), 177-180.

Jentschke, S., Koelsch, S., & Friederici, A. D. (2005). Investigating the relationship of music and language in children: influences of musical training and language impairment. *Ann N Y Acad Sci, 1060*, 231-242.

Jentschke, S., Koelsch, S., Sallat, S., & Friederici, A. D. (2008). Children with specific language impairment also show impairment of music-syntactic processing. *J Cogn Neurosci, 20*(11), 1940-1951.

Knight, R. T., Grabowecky, M. F., & Scabini, D. (1995). Role of human prefrontal cortex in attention control. *Advances in neurology, 66*, 21-36.

Koelsch, S., Gunter, T., Friederici, A. D., & Schroger, E. (2000). Brain indices of music processing: "nonmusicians" are musical. *J Cogn Neurosci, 12*(3), 520-541.

Koelsch, S., Gunter, T. C., Wittfoth, M., & Sammler, D. (2005). Interaction between Syntax Processing in Language and in Music: An ERP Study. *Journal of Cognitive Neuroscience, 17*, 1565-1577.

Koelsch, S., Kasper, E., Sammler, D., Schulze, K., Gunter, T., & Friederici, A. D. (2004). Music, language and meaning: brain signatures of semantic processing. *Nat Neurosci, 7*(3), 302-307.

Koelsch, S., Schmidt, B. H., & Kansok, J. (2002). Effects of musical expertise on the early right anterior negativity: an event-related brain potential study. *Psychophysiology, 39*(5), 657-663.

Koelsch, S., Vuust, P., & Friston, K. (2018). Predictive Processes and the Peculiar Case of Music. *Trends Cogn Sci*.

Kuperberg, G. R. (2007). Neural mechanisms of language comprehension: challenges to syntax. *Brain Res, 1146*, 23-49.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: brain potentials reflect semantic incongruity. *Science, 207*(4427), 203-205.

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

Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological review, 106*(1), 119.

Loui, P., Grent-'t-Jong, T., Torpey, D., & Woldorff, M. (2005). Effects of attention on the neural processing of harmonic syntax in Western music. *Brain Res Cogn Brain Res, 25*(3), 678-687.

Maess, B., Koelsch, S., Gunter, T. C., & Friederici, A. D. (2001). Musical syntax is processed in Broca's area: an MEG study. *Nat Neurosci, 4*(5), 540-545.

Näätänen, R., Gaillard, A. W., & Mäntysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta psychologica, 42*(4), 313-329.

O'Connell, R. G., Dockree, P. M., & Kelly, S. P. (2012). A supramodal accumulation-to-bound signal that determines perceptual decisions in humans. *Nat Neurosci, 15*(12), 1729-1735.

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

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

Przysinda, E., Zeng, T., Maves, K., Arkin, C., & Loui, P. (2017). Jazz musicians reveal role of expectancy in human creativity. *Brain and Cognition, 119*, 45-53.

Roncaglia-Denissen, M. P., Bouwer, F. L., & Honing, H. (2018). Decision Making Strategy and the Simultaneous Processing of Syntactic Dependencies in Language and Music. *Frontiers in Psychology, 9*(38).

Sammler, D., Koelsch, S., Ball, T., Brandt, A., Elger, C. E., Friederici, A. D., . . . Schulze-Bonhage, A. (2009). Overlap of musical and linguistic syntax processing: intracranial ERP evidence. *Ann N Y Acad Sci, 1169*, 494-498.

Sammler, D., Koelsch, S., Ball, T., Brandt, A., Grigutsch, M., Huppertz, H.-J., . . . Schulze-Bonhage, A. (2013). Co-localizing linguistic and musical syntax with intracranial EEG. *NeuroImage, 64*(0), 134-146.

Sammler, D., Koelsch, S., & Friederici, A. D. (2010). Are left fronto-temporal brain areas a prerequisite for normal music-syntactic processing? *Cortex*.

Sammler, D., Koelsch, S., & Friederici, A. D. (2011). Are left fronto-temporal brain areas a prerequisite for normal music-syntactic processing? *Cortex, 47*(6), 659-673.

Slevc, L. R., & Okada, B. M. (2015). Processing structure in language and music: a case for shared reliance on cognitive control. *Psychon Bull Rev, 22*(3), 637-652.

Slevc, L. R., Rosenberg, J. C., & Patel, A. D. (2009). Making psycholinguistics musical: self-paced reading time evidence for shared processing of linguistic and musical syntax. *Psychon Bull Rev, 16*(2), 374-381.

Sun, Y., Lu, X., Ho, H. T., Johnson, B. W., Sammler, D., & Thompson, W. F. (2018). Syntactic processing in music and language: Parallel abnormalities observed in congenital amusia. *Neuroimage Clin, 19*, 640-651.

Treisman, A. M. (1960). Contextual cues in selective listening. *Quarterly Journal of Experimental Psychology, 12*(4), 242-248.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12*(1), 97-136.

van Vugt, M. K., Beulen, M. A., & Taatgen, N. A. (2019). Relation between centro-parietal positivity and diffusion model parameters in both perceptual and memory-based decision making. *Brain Research, 1715*, 1-12.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 June 2019 doi:10.20944/preprints201906.0122.v1

Woldorff, M. G., Gallen, C. C., Hampson, S. A., Hillyard, S. A., Pantev, C., Sobel, D., & Bloom, F. E. (1993). Modulation of early sensory processing in human auditory cortex during auditory selective attention. *Proc Natl Acad Sci U S A, 90*(18), 8722-8726.

Woldorff, M. G., & Hillyard, S. A. (1991). Modulation of early auditory processing during selective listening to rapidly presented tones. *Electroencephalogr Clin Neurophysiol, 79*(3), 170-191.