Music and Language:

Syntactic Interactions in Simultaneous Reading and Production

Joseph C. Clark

University of Rochester

Abstract

It is widely accepted that musical and linguistic syntax share neural processing mechanisms, but there has been little empirical research on whether simultaneous reading and production of music and language results in a similar overlap. In this study, trained Western classical musicians read syntactically ambiguous sentences aloud while simultaneously performing written harmonies or rhythms. Harmonic syntactic interference will be measured in a self-paced reading paradigm while rhythmic syntactic interference will be measured in a post-stimulus-question processing-time paradigm. Expected results demonstrate an interaction between musical and linguistic syntax, lending further support to the Shared Syntactic Resource Integration Hypothesis (SSIRH) and to the dual stream model of language.

Keywords: Syntax, Music, Language, Production, Rhythm

Music and Language:

Syntactic Interactions in Simultaneous Reading and Production

Music is often referred to as “the universal language”. On an intuitive level, music seems able to communicate something emotional and deep across barriers of language and culture. However, how similar are language and music on a structural level, and how similar is their processing on a neural level?

At a first glance, language and music are both hierarchical structures simply built from different units: language uses fundamental units of phonemes, varying across a timbre space (and length, to a limited degree) – while music uses fundamental units of notes varying across dimensions of pitch (melody/harmony) and length (rhythm) (Jackendoff, 2009). In the pitch dimension, a fundamental concept in Western music is that of *key* – that is, which of the twelve pitches is most stable (Slevc, Rosenberg, & Patel, 2003). Hierarchical relationships exist between pitches within a key (the name of the key is known as the *tonic*, the most stable pitch), between *chords* (collections of pitches) within a key, and between keys themselves (Patel, 2003). In particular, it has been noted that the hierarchical relationship of chords within a key is implicitly known by children raised in environments where they are exposed to Western music (Schellenberg, 2005), indicating that these hierarchical relationships are perceptible to all humans, not just those with musical training. However, the pitch dimension does not have any equivalence to linguistic words or semantics, and the harmonic syntactic structure does not encode the same sorts of relationships that linguistic syntactic structure does (lacking, for instance, parts of speech, phrase heads, and means of encoding various dependencies) (Jackendoff, 2009).

In the rhythm dimension, slightly stronger relationships seem to exist between language and music. They share a hierarchical metrical grid: in language, each syllable is a beat in the metrical grid (i.e., the beats are of unequal length) while in music, the grid is isochronous and notes may be longer or shorter than a beat (i.e., the notes are of unequal length) (Jackendoff, 2009). Musical rhythm, however, contains a deep recursive hierarchy while linguistic rhythm does not (the closest linguistic analogue being intonational phrasing, which of necessity is bounded by the breath); contrarily, an intonational phrase has internal constituents (e.g., phonological words) which musical rhythmic structures do not exhibit (Jackendoff, 2009). It is interesting, however, to note that the variability of vowel length in a composer's native language positively correlates with the variability of rhythm in their musical themes (Patel & Daniele, 2003). This lends credence to the theory that music and language have overlapping neural correlates.

Literature Review

Music and Phonology

Despite the fact that music has no direct equivalent of linguistic phonology, interactions have been found between the two. Participants are able to more quickly identify phonemes when sung on a stable tonic chord (that is, the chord built on the tonic pitch) than when sung on a harmonically congruent but less stable chord (Bigand, Tillmann, Poulin, D'Adamo, & Madurell, 2001). Since participants were not asked to attend to the harmonic structure, this indicates that harmonic relationships – indeed, even identification of key, which requires attention to the entire progression rather than just the final chord – occur subconsciously and automatically. Musicians and non-musicians exhibited no difference in performance, suggesting that this is not a result of extensive musical training.

Drawing from fMRI data, overlapping brain areas are activated when exposed to speech, vocalise (textless music), and song (texted music), particularly in the M/STG and I/MFG (Schön et al., 2010). Prior research implicates that language activates these areas in the left hemisphere while music activates them on the right; while the fMRI data supports this conclusion, bilateral activation occurred for all conditions to varying degrees. This suggests an overlap of phonological/lexical processing and melodic processing.

A study by Kolinsky, Lidji, Peretz, Besson, and Morais (2009) provides an important caveat to the above results: by measuring reaction times, they conclude that melodies interfere with the processing of vowels but not consonants. It thus appears that, while phonological and melodic processing have some degree of overlap, their neural correlates are not identical.

Music and Semantics

There is far less empirical support for an interaction between musical syntax and linguistic semantics. In a study by Bonnel, Faita, Peretz, & Besson (2001), participants listened to modified excerpts from French operatic songs where the last note was either in-key or out-of-key and the last word was either semantically congruous or incongruous. They were asked to identify melodic incongruities, semantic incongruities, or both. Given that there was no difference in performance between the tasks, and no correlation between the different cases of double-identification, the authors conclude that musical syntactic processing and linguistic semantic processing do not interact. Once again, this result held up identically between musicians and non-musicians.

ERP data corroborates this conclusion: Listening to out-of-key chords does not affect the processing of semantically improbable (but syntactically correct) sentences (Koelsch, Gunter, Wittofth, & Sammler, 2005). A semantic control in Slevc, et al. (2003) (discussed further under

Music and Syntax, below) also indicates no relationship between harmonic syntax and linguistic semantics.

It is worth underlining that these results are not surprising: given that music has no clear equivalent to semanticity, there is no proposed mechanism whereby musical syntactic processing would interfere with linguistic semantic processing. That said, data also exists that supports a relationship between these two processes. Using a similar experimental design to Bigand et al. (2001), participants are able to more quickly make a word/nonword decision for semantically related words when sung on a stable tonic chord than when sung on a harmonically congruent but less stable chord (Poulin-Charronnat, Bigand, Madurell, & Peereman, 2005). However, the data suggests that unexpected chords divert attentional resources from linguistic processing. This explanation of attention modulation thus does not contradict the earlier findings of a lack of overlap in music syntactic and linguistic semantic processing networks.

Music and Syntax

Syntax is the point of clearest overlap between music and language. Although the exact nature of the hierarchical structure differs, the fact remains that similar syntactic calculations must take place in both domains. Thus, it seems reasonable to postulate that there is some neural overlap in the syntactic processing of both domains. However, some neuropsychological data questions this reasonable assumption, since there are documented instances of amusia (difficulty perceiving harmonic syntactic relationships) without aphasia (Peretz, 1993) and aphasia without amusia (Luria, Tsvetkova, & Futer, 1965). On the other hand, the P600, an ERP associated with linguistic syntactic processing, is also found in musicians listening to harmonic progressions (Patel, Gibson, Ratner, Besson, & Holcomb, 1998). MEG (Maess, Koelsch, Gunter, & Friederici, 2001) and fMRI (Tillmann, Janata, & Bharucha, 2003) studies corroborate that there is a relationship between harmonic and linguistic syntactic processing, even going so far as to suggest that the classic language areas – Broca's Area and Wernicke's Area – are implicated in harmonic syntactic processing (Koelsch, et al., 2002) (see also Patel, 2003 for further references).

To make sense of this conflicting data, Patel (2003) suggests there's a meaningful (and physical) distinction between syntactic representation (i.e., domain-specific syntactic rules) and syntactic processing (i.e., a workspace where those rules are applied to input). Both extant language processing models (cf. Gibson's Dependency Locality Theory (Gibson, 1998)) and music processing models (cf. Lerdahl's Tonal Pitch Space theory (Lerdahl, 2001)) propose a mechanism that mentally integrates each incoming stimulus (be it word or harmony) with the previously established structure. They additionally both propose that integration is influenced by the abstract distance of the incoming stimulus from previous stimuli. Thus, Patel (2003) proposes the Shared Syntactic Integration Resource Hypothesis (SSIRH), holding that linguistic and musical syntax are processed using the same neural networks while their respective representations remain neutrally distinct. The SSIRH thus explains the apparent conflict in the above studies: studies supporting the overlap of harmonic and linguistic syntactic processing are a result of the shared integration region, while those supporting a dissociation are a result of damage to either the linguistic or musical syntactic representation regions.

The SSIRH predicts that individuals with Broca's aphasia (difficulty with syntactic integration of language) should also exhibit difficulty with processing harmonic syntactic relationships, since it posits that syntactic processing is shared between the two domains (Patel 2003). This prediction is borne out (Patel, Iversen, Wassenaar, & Hagoort, 2008). While the performance of participants with Broca's aphasia is closer to that of controls on harmonic syntactic tasks than linguistic syntactic tasks, it is still significantly lower. This is true even controlling for potential confounds such as pitch perception and auditory short-term memory. However, it is unclear why the effect is less for harmonic syntax than linguistic syntax.

ERP data also supports the SSIRH. Canonically, an ELAN (Early Left Anterior Negativity) indicates initial linguistic syntactic processing while an ERAN (Early Right Anterior Negativity) indicates initial musical syntactic processing (Koelsch et al., 2005), suggesting that the processing of linguistic and musical syntax is similar and occurs simultaneously. Additionally, the ELAN was significantly less pronounced when an out-of-key chord occurred simultaneously, that could not be accounted for by an otherwise physically deviant tone (in loudness or timbre).

Further support for the SSIRH comes from an experiment investigating simultaneous linguistic and musical input, in the form of a sung sentence (Fedorenko, Patel, Casasanto, Winawer, & Gibson, 2009). The authors manipulated linguistic syntactic complexity (subject-extracted vs. object-extracted relative clauses) and musical complexity (in-key or out-of-key critical note), as well as an auditory-anomaly control (suddenly louder critical note) that did not affect processing. After each sentence, a YES/NO question was asked; the difference in accuracy on these questions for subject-extracted vs. object-extracted (more difficult) relative clauses significantly increased when the critical note was out-of-key (but was not affected by the auditory-anomaly control). This suggests that increased musical syntactic processing can interfere with linguistic syntactic processing at critical moments.

A similar experiment (Slevc, et al., 2003) uses a different experimental paradigm to reach the same conclusions. Similarly to Fedorenko, et al. (2009), the study manipulates both linguistic syntactic complexity (in the form of ambiguous “garden-path” sentences, which often require syntactic reanalysis midway through, rather than relative clauses) and musical complexity (in-key or out-of-key critical chord, rather than note). Rather than testing comprehension, the authors instead use a self-paced reading paradigm, aurally presenting a chord with each word or short phrase. They find that the critical word reading time is substantially longer when it is presented with an out-of-key chord, while there was no interaction with unexpected semantic and timbral controls.

However, all of these studies focus solely on the processing of musical harmonic syntax. It appears that research has not yet been done on the interaction of musical harmonic and linguistic syntax in reading and production. Based on the recent dual-stream model of language (Hickok & Poeppel, 2007), it seems likely that processing and production of language utilize the same brain regions in different orders. Thus, these interactions between linguistic and musical syntax observed in the above studies ought to be borne out in a study of simultaneous linguistic and musical production.

It also appears that research has not yet been done on the interaction of musical rhythmic and linguistic syntax in either processing or production. This is certainly more difficult to measure directly, given the explicitly temporal nature of rhythm, but results in this area can either substantiate or discredit Jackendoff (2009)'s claim of the metrical similarities between music and language.

Methods

Because of the complexity of any task involving simultaneous production of music and speech, participants will be limited to trained performers of Western classical music. Several studies (Bigend, et al., 2001; Bonnel, et al., 2001, Slevc, et al., 2003) show that musicians and non-musicians do not differ in their processing tendencies. Thus, the sole use of trained musicians should not affect the data, while ensuring the ability to read music and eliminating any differential task-complexity effects on processing. Additionally, the requirement of producing harmonies while simultaneously speaking limits potential subjects to those who perform on manually-operated harmonic instruments (piano, guitar, organ, etc.).

Experiment 1 will investigate the relationship between harmonic syntactic production and linguistic syntactic production. Closely following the method of Slevc, et al. (2001), participants will perform written harmonic progressions while speaking sentences printed as lyrics with the given progression. Half the sentences will be syntactically unambiguous and half will be syntactically ambiguous (“garden path” sentences). Likewise, half the harmonic progressions will contain an in-key chord at the point of disambiguation while half will contain an out-of-key chord. By using the same materials, it is ensured that any interference is syntactic in nature (since Slevc et al. (2001) demonstrated that non-syntactic incongruities such as timbre did not interfere with linguistic syntactic processing; this is corroborated in Fedorenko (2009)), and thus likely a result of shared syntactic processing resources. This experiment will use a similar self-paced reading paradigm, where the screen advances to the next harmony/word on the press of a foot pedal (since the hands are occupied). The reading time of the disambiguation point (and surrounding words) will be measured in all four conditions.

Experiment 2 will investigate the relationship between rhythmic syntactic production and linguistic syntactic production. Following the method of Fedorenko, et al. (2009) (since it likewise used temporally fixed stimuli), participants will tap printed rhythms while speaking sentences, with four conditions as in experiment 1. Printed rhythms will be of equivalent complexity and equal duration (one beat in duple meter, at a tempo of one beat per 1000ms), with the exception of the out-of-meter rhythm at the point of disambiguation (which will last one and a half beats in duple meter). One word will be presented with each rhythm, and the slides will automatically advance at 600ms after onset to mimic the "looking-ahead" done by trained musicians (Madell & Hébert, 2008). Since reading time cannot be measured in this paradigm, the processing time and response rate accuracy of YES/NO comprehension questions following each target sentence will be measured.

Results/Discussion

Experiment 1

Given the wealth of previous literature supporting the SSIRH, it is likely that Experiment 1 will find slower processing times for syntactically ambiguous sentences when paired with an out-of-key chord. (It is likely that slower processing times for syntactically unambiguous sentences will also be found, although this effect is likely to be less significant, since fewer syntactic processing resources are required in this case.) These results would support the SSIRH, as they indicate a neural overlap of linguistic and musical syntactic production, and the dual-stream model, as they allow for the same syntactic integration region to be used in both processing and production. Negative results (i.e., no difference in processing time across harmonic conditions) may indicate issues with the dual-stream model (e.g., the syntactic integration region used in processing is not the same one used in production). Task complexity is certainly worth keeping in mind as a possible confound, although it ought to affect all conditions equally. (That said, it is likely that self-paced reading times will be significantly longer than those found in other studies given this complexity.)

Experiment 2

While there appears to be no prior research on the interaction of rhythmic and linguistic syntactic processing or production, the SSIRH would predict that there would be an interactive effect, at least at the level of processing. If Experiment 1 yields positive results, indicating a shared syntactic integration region for processing and production, then Experiment 2 will likely also yield positive results (i.e., slower processing times for syntactically ambiguous sentences paired with an out-of-meter rhythm, with a lesser effect on syntactically un-ambiguous sentences). Likewise, if Experiment 1 yields negative results, indicating linguistic and harmonic syntax are produced separately, it is likely that Experiment 2 will also yield negative results (else this would indicate that linguistic and rhythmic syntax have production overlap while linguistic and harmonic syntax do not). Possible confounds include task complexity (which again should affect all four conditions equally) and the disorientating effect that an out-of-meter rhythm has on subsequent stimuli (this effect is present but not as strong with out-of-key chords). If this is in fact a confound, it should affect both syntactic conditions equally, and can be investigated in a control study by placing out-of-meter rhythms at various points in the sentence (both syntactically critical and non-critical) and observing if there is a difference in effect.

General Discussion

Mixed results (i.e., positive results for Experiment 1 AND negative results for Experiment 2 and vice versa) are thus not expected, and likely indicate issues with experimental design. Given the apparent novelty of research in both simultaneous production and the interaction of rhythmic and linguistic syntax, further research is warranted along both avenues. Simultaneous production studies modelled after the various simultaneous processing studies (both in terms of experimental method and interaction of musical syntax with linguistic phonology and linguistic semantics) would provide a more complete picture. The investigation of the interaction of rhythmic and linguistic syntax processing would fill a gap in the literature, particularly if Experiment 2 yields negative results.

Conclusion

While the interaction of musical and linguistic syntax in production may have profound theoretical implications for the SSIRH and the dual-stream model, it even more fundamentally deals with the universal human behavior of singing. Positive results in either experiment have strong implications in the field of music education, whether it be designing pedagogy to deal with anticipated places of difficulty in repertoire (where both musical and syntactic anomalies align) or considerations in composing choral works for younger ensembles. Understanding how we produce song as humans brings us a little closer to understanding ourselves.

References

Bigand, E., Tillmann, B., Poulin, B., D'Adamo, D.A., & Madurell, F. (2001). The effect of harmonic context on phoneme monitoring in vocal music. *Cognition, 81,* B11-B20. Retrieved from www.elsevier.com/locate/cognit

Bonnel, A., Faita, F., Peretz, I., & Besson, M. (2001). Divided attention between lyrics and tunes of operatic songs: Evidence for independent processing. *Perception & Psychophysics, 63(7)*, 1201-1213.

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. doi:10.3758/MC.37.1.1

Gibson, E. (1998). Linguistic complexity: locality of syntactic dependencies. *Cognition, 68,* 1-76.

Hickok, G. & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Review Neuroscience, 8*, 393-402. doi:10.1038/nrn2113

Jackendoff, R. (2009). Parallels and nonparallels between language and music. *Music Perception: An Interdisciplinary Journal, 26(3),* 195-204. doi: 10.1525/mp.2009.26.3.195

Koelsch, S., Gunter, T.C., Cramon, D.Y., Zysset, S., Lohmann, G., Friderici, A.D. (2002). Bach speaks: a cortical "language-network" serves the processing of music. *Neuroimage, 17,* 956-966.

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(10),* 1565-1577.

Kolinsky, R., Lidji, P., Peretz, I., Besson, M., & Morais, J. (2009). Processing interactions between phonology and melody: Vowels sing but consonants speak. *Cognition, 112,* 1-20. doi:10.1016/j.cognition.2009.02.014

Lerdahl. F. (2001). *Tonal pitch space*. Oxford: Oxford University Press.

Luria, A., Tsvetkova, L., & Futer, J. (1965). Aphasia in a composer. *Journal of Neurological Science, 2,* 288-292.

Madell, J., & Hébert, S. (2008). Eye movements and music reading: Where do we look next? *Music Perception: An Interdisciplinary Journal, 26(2),* 157-170. DOI:10.1525/MP.2008.26.2.157

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

Patel, A.D. (2003). Language, music, syntax and the brain. *Nature Neuroscience, 6(7)*, 674-681.

Patel, A.D., & Daniele, J.R. (2003). An empirical comparison of rhythm in language and music. *Cognition, 87*, B35-B45. doi:10.1016/S0010-0277(02)00187-7

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.

Patel, A.D., Iversen, J.R., Wassenaar, M., & Hagoort, P. (2008). Musical syntactic processing in agrammatic Broca's aphasia. *Aphasiology, 22(7-8),* 776-798. doi: 10.1080/02687030701803804

Peretz, I. (1993). Auditory atonalia for melodies. *Cognitive Neuropsychology, 10,* 21-56.

Poulin-Charronnat, B., Bigand, E., Madurell, F., & Peereman, R. (2005). *Cognition, 94*, B67-B78. doi:10.1016/j.cognition.2004.05.003

Schellenberg, E.G., Bigand, E., Poulin-Charronnat, B., Garnier, C., & Stevens, C. (2005). Children's implicit knowledge of harmony in Western music. *Developmental Science, 8(6)*, 551-566.

Schön, D., Gordon, R., Campagne, A., Magne, C., Astésano, C., Anton, J., & Besson, M. (2010). Similar cerebral networks in language, music and song perception. *NeuroImage, 51,* 450-461. doi:10.1016/j.neuroimage.2010.02.023

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. *Psychonomic Bulletin & Review , 16(2),* 374-381. doi:10.3758/16.2.374

Tillmann, B., Janata, P., & Bharucha, J.J. (2003). Activation of the inferior frontal cortex in musical priming. *Cognitive Brain Research, 16,* 145-161.