



## The language of music: Common neural codes for structured sequences in music and natural language

Jeffrey N. Chiang<sup>a,1</sup>, Matthew H. Rosenberg<sup>a,1</sup>, Carolyn A. Bufford<sup>a</sup>, Daniel Stephens<sup>b</sup>, Antonio Lysy<sup>b</sup>, Martin M. Monti<sup>a,\*</sup>

<sup>a</sup> Department of Psychology, University of California Los Angeles, Los Angeles, CA, USA

<sup>b</sup> Department of Music, UCLA Herb Alpert School of Music, University of California Los Angeles, Los Angeles, CA, USA

### ARTICLE INFO

#### Keywords:

Language  
Music  
Neuroimaging  
Structured sequences  
Multivariate Pattern Analysis (MVPA)

### ABSTRACT

The ability to process structured sequences is a central feature of natural language but also characterizes many other domains of human cognition. In this fMRI study, we measured brain metabolic response in musicians as they generated structured and non-structured sequences in language and music. We employed a univariate and multivariate cross-classification approach to provide evidence that a common neural code underlies the production of structured sequences across the two domains. Crucially, the common substrate includes Broca's area, a region well known for processing structured sequences in language. These findings have several implications. First, they directly support the hypothesis that language and music share syntactic integration mechanisms. Second, they show that Broca's area is capable of operating supramodally across these two domains. Finally, these results dismiss the recent hypothesis that domain general processes of neighboring neural substrates explain the previously observed “overlap” between neuroimaging activations across the two domains.

### 1. Introduction

A central intuition in the study of human language as a cognitive phenomenon is the idea that, while listening to a linear signal such as speech, our minds spontaneously build abstract and structured hypotheses representing how discrete elements within a sequence relate to each other (Chomsky, 1957, 1965; Fitch and Martins, 2014; Jackendoff, 2002; Lashley, 1951; Monti, 2017). The use of such representations is most clearly displayed in natural language (Berwick, Friederici, Chomsky, & Bolhuis, 2013; Ding, Melloni, Zhang, Tian, & Poeppel, 2015), but also characterizes other aspects of human cognition, such as logical reasoning (Monti & Osherson, 2012; Osherson, 1975), algebraic cognition (Maruyama, Pallier, Jobert, Sigman, & Dehaene, 2012; Monti, Parsons, & Osherson, 2012; Varley, Klessinger, Romanowski, & Siegal, 2005), and music cognition (Katz & Pesetsky, 2011; Lerdahl, 2001; Patel, 2003), among others. The relationship between the syntactic operation of language and the syntax-like operations of other aspects of human cognition has thus been at the center of a long-standing debate concerning the degree to which human thought is embedded within, or enabled by, natural language (e.g., Lashley, 1951; Boeckx, 2010; Gleitman & Papafragou, 2013; Fitch & Martins, 2014; Fitch 2014;

Monti, 2017).

Lashley (1951) commented on the prevalence of structured sequences across domains, noticing that they exhibited the following three properties: (1) connectedness; i.e. no node is isolated from the others, (2) a root element; i.e. “sentence” or “chord” that is superior to others and (3) acyclic structure; establishing order as a unique property (Fitch and Martins 2014; Lashley 1951). In the context of music cognition, the analogy with the structural aspects of language is particularly pronounced. As discussed elsewhere (e.g., Lerdahl & Jackendoff, 1985; Patel, 2003; Fadiga, Craighero, & D'Ausilio, 2009; Fitch 2014; Peretz, Vuvan, Lagrois, & Armony, 2015), music and language are both characterized by discrete elements (e.g., words, chords) which can be (recursively) combined, according to specific rules, to form organized structures (e.g., sentences, melodies) which are typically encoded within linear, time-dependent, signals.

Nonetheless, whether this analogy is substantial or merely superficial remains a debated issue (cf., Peretz et al., 2015). At one end of the spectrum, it has been proposed that language and music are governed by the very same syntactic processes applied to different building blocks (e.g., words vs. notes). According to this view, “[a]ll formal differences between language and music are a consequence of

\* Corresponding author at: 7461E Franz Hall, Department of Psychology, University of California Los Angeles, Los Angeles, CA 90095, USA.

E-mail address: [monti@ucla.edu](mailto:monti@ucla.edu) (M.M. Monti).

<sup>1</sup> Equal contribution.

differences in their fundamental building blocks; in all other respects, language and music are identical” (Katz & Pesetsky, 2011). Along similar lines, it has been proposed that the common representations underlying the structure processing in language and music can be localized to the neural mechanisms encapsulated within the left inferior frontal gyrus (IFG; often referred to as Broca’s Area), a region hypothesized to operate as a “supramodal hierarchical parser” (Fadiga et al., 2009; Tettamanti & Weniger, 2006). Consistent with this view, a rapidly growing neuroimaging literature has shown music processing to recruit cortical regions overlapping with areas known to be involved in syntactic and semantic aspects of natural language processing (Patel, Gibson, Ratner, Besson, & Holcomb, 1998; Maess, Koelsch, Gunter, & Friederici, 2001; Koelsch et al., 2002; Tillmann, Janata, & Bharucha, 2003; Koelsch et al., 2004; Koelsch, Fritz, Schulze, Alsop, & Schlaug, 2005; Brown, Martinez, & Parsons, 2006; see Rogalsky, Rong, Saberi, & Hickok, 2011, for a conflicting result). Nonetheless, while the observation of overlapping neural substrates is often taken to imply the presence of shared neurocognitive representations between language and music, this is not necessarily the case (Peretz et al., 2015) and indeed has never been shown to be true. This “missing link” in the neuroscientific literature leaves open the possibility that commonly recruited areas of the brain might, in fact, represent very different operations that do not translate, or align, across the two domains, or that are entirely unrelated to the processing of these relationships. In line with this observation, it has been suggested that language and music are in fact better thought of as modular and largely independent of each other (Marin & Perry, 1999; Peretz & Coltheart, 2003). In support of this view, a rich neuropsychological literature has described cases of individuals who exhibit amusia in the absence of aphasia, as well as aphasia in the absence of amusia (Luria et al., 1965; Peretz, 1993; Peretz et al., 1994; Ayotte, Peretz, Rousseau, Bard, & Bojanowski, 2000; Piccirilli, Sciarma, & Luzzi, 2000; Ayotte, Peretz, & Hyde, 2002).

The reason for the contradicting evidence is still a matter of debate. According to some, the fracture between neuropsychological and neuroimaging findings can be reconciled with a middle-ground solution in which language and music are viewed as partially overlapping systems (Patel, 2003; Patel, Iversen, Wassenaar, & Hagoort, 2008). Under this view, referred to as the shared syntactic integration resource hypothesis, language and music are characterized by both domain-specific (i.e., separate) and shared processes. The domain-specific processes relate to the particular features of each syntax, which are recognized as architecturally different, while shared processes provide neural resources for the activation of the relevant stored syntactic representations (Patel, 2012). According to others, the inconsistency between the two sets of findings might instead be due to experimental and neuroanatomical considerations (Fedorenko & Varley, 2016). Specifically, the overlap often reported, in neuroimaging studies, in left inferior frontal regions could be a reflection of task-general demands tied to the use of structural-violation paradigms (e.g., the P600 and the early left/right anterior negativity effects reported in electrophysiological studies; Janata, 1995; Maess et al., 2001; Koelsch et al., 2002, 2005; Steinbeis and Koelsch, 2008; Tillmann et al., 2003; and later localized to the inferior frontal gyri through neuroimaging; Musso et al., 2015; Kunert, Willems, Casasanto, Patel, & Hagoort, 2015). Deviant events are indeed likely to elicit ancillary processes including attentional capture, detection of violated expectations, or error correction, regardless of whether the violation applies to natural language, music, arithmetic, or motor sequences. Such processes are unrelated to the extracting or forging of structured sequences and are known to elicit activation in domain-general regions (proximal or partially overlapping with Broca’s Area; see Fedorenko & Varley, 2016, for a detailed discussion).

In the present study, we address the relationship between the mechanisms of natural language and those of music in a 3 Tesla functional magnetic resonance imaging (fMRI) within-subjects design in which competent musicians generate structures in language (active/passive voice sentences versus repeating a verb) and music (root/second-

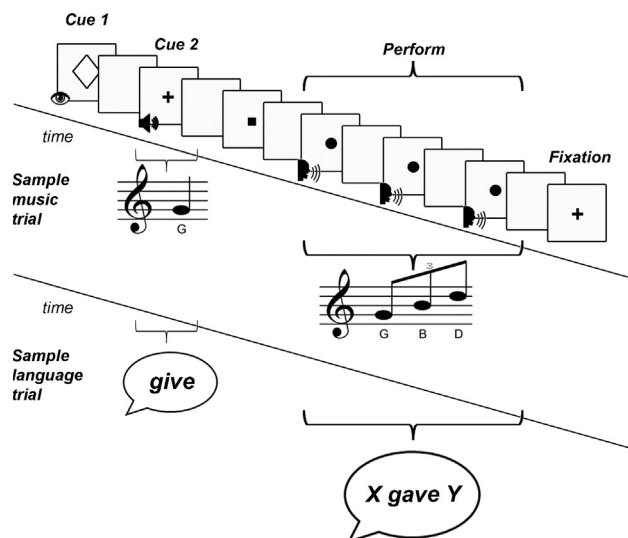


Fig. 1. Experimental design. Sample music and language trials timelines.

Table 1  
Sample cues and stimuli.

Language	Structure Active	Structure Passive	Non-Structure Repeat
Cue 1	♦	♣	○
Cue 2	“Pay”	“Give”	“Tell”
Response	“X paid Y”	“X was given Y”	“Tell, Tell, Tell”
Music	Root Position	IInd Inversion	Repeat
Cue 1	♦	♣	○
Cue 2	“C”	“D”	“E”
Response	“C-E-G”	“D-G-B”	“E-E-E”

inversion position ascending triads versus repeating a note; cf., Fig. 1 and Table 1). Crucially, we employ a (rarely explored) generation task to avoid the confound of salient events, and we use a multivariate cross-classification approach to resolve the interpretational ambiguity present in the previous neuroimaging literature (which has been specifically advocated for; see Peretz et al., 2015), thereby helping resolve the question of whether natural language and music share a common underlying neural code for representing structured sequences.

## 2. Methods

### 2.1. Participants

We recruited 21 total participants to reach the predetermined sample size (N = 20, 8 female participants) based on previous literature (Musso et al., 2015: N = 11; Kunert et al., 2015: N = 19; Koelsch et al., 2002: N = 20). An additional subject was recruited because the data from one of the participants exhibited excessive motion during the procedure (see below). Participants received \$50 compensation for taking part in the experiment. All participants were native English speakers, right handed, and competent musicians currently enrolled in the UCLA Herb Alpert School of Music. Participants were only enrolled if they could demonstrate proficiency in singing/generating both a root position and IInd inversion ascending triad arpeggio. Participants with perfect pitch were excluded. Participants signed informed consent prior to taking part in the session, as per the procedures approved by the UCLA Institutional Review Board.

## 2.2. Stimuli

For both materials (i.e., “language” and “music” trials), the first cue was delivered visually, by presenting one of three icons in the middle of the screen. A ‘◇’ symbol indicated an active or root position trial (depending on whether the second cue was a word or a note, respectively); a ‘♣’ symbol indicated a passive or a IInd inversion trial; a ‘∪’ symbol indicated a non-structured (i.e., repeat) trial. The second cue was delivered aurally and consisted of either a verb or a note, thus revealing whether the trial was a language or music trial, and allowing disambiguation of the instruction provided by the first cue. (See Table 1 and Fig. 1 for sample stimuli.) Cues for language trials consisted of seven monosyllabic, reversible, present tense verbs (i.e., “bring,” “tell,” “teach,” “throw,” “leave,” “give,” “pay”). Cues for the music trials consisted of 7 notes recorded with an electronic piano. The notes comprised a chromatic scale (i.e. each consecutive pitch was separated from the following by a semitone, or a half step). Participants were allowed to choose, among three sets of musical cues (high, middle, or low), whichever pitch range most comfortably matched their vocal range. For each set, a perfect fifth (7 semitones/half steps) separated the lowest and highest pitches of the cues. Combining the 7 cues (i.e., words/notes), 3 generative rules (2 structured, 1 repeat), and 2 materials (i.e., language, music), resulted in 42 unique trials (i.e., 21 per material type).

## 2.3. Experimental design

Each participant viewed the 42 unique trials twice (once in the first two runs, once in the second two runs). Trial types were equally distributed across 4 runs, and, within each, randomly presented. Stimuli were presented using PsychoPy (Peirce, 2008); visual cues were displayed through a custom-made MRI-compatible projection system while auditory cues were delivered through a Magnetic Resonance headphone system. As shown in Fig. 1, each trial started with the generative rule cue (i.e., ‘◇’, ‘♣’, ‘∪’), displayed on screen for 1.5 s, followed by the second cue (i.e., word or note) presented, aurally, for 1.8 s. After a variable jitter (between 6 and 8 s), a fixation symbol blinked four times (with a cycle of 0.8 s of display and 0.35 s interval). The first blink (with a black square symbol) served as a warning that the “performance/response” period was to begin. The following three blinks (with a black circle symbol) marked the performance/response period and provided a tempo for responding. The tempo was never varied, neither within nor across subjects, and was only employed to provide participants with a consistent rhythm for responding.

Finally, a variable length fixation screen (with a random jitter between 5 and 7 s chosen, on a trial-by-trial basis, from an exponential distribution) separated each trial from the subsequent one. Each run lasted, on average, 293.57 s (S.D. = 15.81). Participants were trained to asymptotic performance prior to the imaging session, in a separate room, after having signed informed consent. The experimenter corrected any errors the participant made until satisfactory performance was achieved (less than 2 errors per block of trials). Training ceased when participants could perform at least 12 out of 13 trials correctly, minimizing the sound production time across conditions.

## 2.4. Data acquisition

Data were acquired on a 3 Tesla Siemens Tim Trio Magnetic Resonance Imaging (MRI) scanner at the One Mind Center for Cognitive Neuroscience at UCLA. Structural data were acquired using a T1-weighted sequence (MP RAGE, TR = 1900 ms, TE = 2.26 ms, voxel size 1 mm<sup>3</sup> isovoxel). Blood oxygenation level dependent (BOLD) data were acquired with a T2\*-weighted Gradient Recall Echo sequence (TR = 3000 ms, TE = 35 ms, 45 interleaved slices, voxel size 3 × 3 × 3.3 mm) with prospective motion correction in order to reduce the impact of subject motion during performance.

## 2.5. Data preprocessing

Data analysis was carried out using FSL (Smith et al., 2004). Prior to analysis, data underwent a series of conventional preprocessing steps including motion correction, slice-timing correction (using Fourier-space time-series phase-shifting), spatial smoothing using a Gaussian kernel of 5 mm full-width half-max, and highpass temporal filtering (Gaussian-weighted least-squares straight line fitting, with  $\sigma = 50.0$  s). Data from each individual run were analyzed employing a univariate general linear model approach (Monti, 2011) inclusive of a pre-whitening correction for autocorrelation. Following current convention, any participant exhibiting average motion greater than 3 mm was excluded (N = 1).

## 2.6. Univariate analysis

For each run of each participant, a univariate analysis was conducted using, as the main variables of interest, 6 regressors, one per trial type (i.e., language active voice, language passive voice, language repeat, music root position, music IInd inversion position, music repeat). Regressors marked the performance/response period of each trial (see Fig. 1). A number of additional nuisance regressors were employed to model cue periods, motion (including first and second derivatives, and their difference), as well as the short intervals between the second cue and task performance. This last regressor is particularly important since it parcels out periods in which subjects are likely to be engaging in strategies in anticipation of the task, which, in the absence of any participant feedback, are un-controlled and thus difficult to interpret. For each run we computed 4 contrasts: structured versus repeat trials for language and music materials (“simple effect” contrasts), separately, and the interaction between the two simple effects (“interaction contrasts”) in both directions (i.e., simple effect of structured trials in language greater than the simple effect of structured trials in music, and vice versa). Data from individual runs were aggregated employing a mixed effects model (i.e., employing both the within- and between-subject variance), and using automatic outlier detection. Z (Gaussianised T) statistic images were thresholded using a cluster correction of  $Z > 2.3$  and a (corrected) cluster significance threshold of  $P = 0.05$ .

## 2.7. Multivariate analysis

The input to the multivariate analysis was a set of volumes of regression coefficients (i.e., “ $\hat{\beta}$ ”) marking the magnitude of activation, for each voxel, in each trial (per subject). These trial-wise “patterns of activations” were obtained by employing the iterative Least Squares – Separate approach (LS-S; Mumford, Turner, Ashby, & Poldrack, 2012) in which a separate GLM is run (here, using FILM with local autocorrelation) for each trial. At each iteration, one regressor marks the trial of interest, while all remaining trials are collapsed into a nuisance regressor (see Mumford et al., 2012, Fig. 1 for a visual depiction of this approach). This approach has been shown, in simulations, to produce activation estimates that have the highest correlation with true activation magnitudes (Mumford et al., 2012), and has also been shown to adapt best to multivariate analyses when used in conjunction with full randomization of trials (different for each subject) and with equal inter-stimulus interval across condition (Mumford, Davis, & Poldrack, 2014), as we have done. The patterns of activation were then concatenated across time to construct a subjectwise “ $\hat{\beta}$ -series” of activation magnitude per trial per voxel (Rissman, Gazzaley, & D’Esposito, 2004).

In order to assess whether natural language and music share underlying neural representations, we employed a cross-classification searchlight analysis using a linear support vector machine (SVM) algorithm. Cross-classification was performed by training the SVM classifier to recognize structure vs. repeat trials in one domain, and then attempting to classify structure vs. repeat structure trials in the other

domain (“M2L” and “L2M” cross-classifications for training on music and testing on language and vice versa, respectively). L2M cross-classifications were performed over voxels found significant in the structure minus repeat trials for language materials (only); M2L classifications were performed over voxels found significant in the structure minus repeat trials for music materials (only). Because the significant voxels in the two univariate contrasts could overlap, this feature selection ensures that the training and testing datasets for each type of cross-classification (i.e., L2M, M2L) remain completely separate, thereby avoiding any bias in the analysis. Classifications were performed on a single subject basis, in native space, employing a 6 mm radius searchlight approach (Kriegeskorte, Goebel, & Bandettini, 2006). To account for the imbalance between the number of structure and repeat trials (28 and 14, respectively, per each domain) and avoid biasing the classifier, we performed a resampling procedure in which, at each of 1000 iterations, a subsample of 14 (structured) trials was randomly selected, in order to train and test the classifier on a matching number of trials across conditions. Results across the 1000 iterations were averaged to yield a single classification accuracy value for each searchlight sphere.

Then, in order to assess whether the two structured sequences of each domain could be distinguished from one another on the basis of patterns of brain activity within Broca’s area (i.e., Language Active Voice ( $L_{AV}$ ) versus Language Passive Voice ( $L_{PV}$ ) and Music Root Position ( $M_{RP}$ ) versus Music 2nd Inversion ( $M_{SI}$ )), we performed a second searchlight analysis (Kriegeskorte et al., 2006). Specifically, for each domain separately, we performed an SVM classification on a single subject basis, in native space, employing a 6 mm radius searchlight. At each iteration of the searchlight, classification accuracy was assessed using a leave-one-run-out cross validation procedure in which the SVM was trained on the trials from 3 runs, and accuracy was assessed on the trials from the remaining run. This procedure was repeated for each run, and the 4 accuracies were then averaged to obtain overall accuracy for each classification.

For both classification analyses, statistical significance was assessed, at the group level, employing a permutation-based sign test and against a criterion of  $p = 0.05$  corrected for multiple comparisons at the cluster level (using FSL’s threshold free cluster enhancement, Winkler, Ridgway, Webster, Smith, & Nichols, 2014). At the single subject level, significance was assessed with a permutation procedure, in which the classification was repeated 1000 times, with shuffled testing labels, to construct a null distribution for each voxel (cf., Etzel & Braver, 2013). Classifications falling within the top 5% of the null distribution were considered significant.

### 3. Results

#### 3.1. Univariate analysis

The simple effect contrast of structure versus repeat trials for language materials uncovered a set of expected activations in left inferior frontal gyrus (including its *pars opercularis* and *triangularis*, in Brodmann Areas [BA] 44, 45), posterior middle and superior temporal cortices (BA 21, 22), bilateral parietal (spanning BA 7, 40) and medial (BA 6), middle (BA 8), and superior frontal (BA 6) areas (mostly left lateralized; see Fig. 2, below, and Table S1 in the Supplemental Material available online for complete list of local maxima).

Supplementary data associated with this article can be found, in the online version, at <https://doi.org/10.1016/j.bandl.2018.07.003>.

When performed on music trials, the same contrast uncovered a number of activation clusters across bilateral frontal and parietal regions (see Fig. 2 and Table S2). The frontal cluster included bilateral maxima in the inferior frontal gyri (spanning its *pars opercularis* in BA 44, *triangularis* in BA 45, and *orbitalis* in BA 47), rostral insular cortex (spanning BA 13 and its junction with 45 and 47), as well as bilateral foci across middle (BA 6), superior (BA 6, 8) frontal, and cingulate (BA 32) gyri. In addition, bilateral activations were observed in the inferior

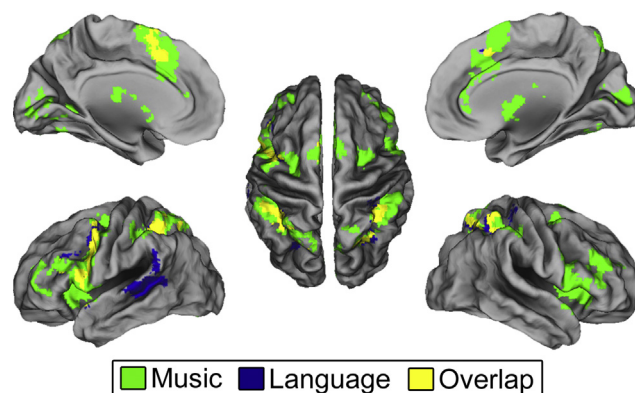


Fig. 2. Univariate result. Overlay of the structure versus repeat contrast results for language (blue) and music (green) (yellow marks overlap between the two tasks).

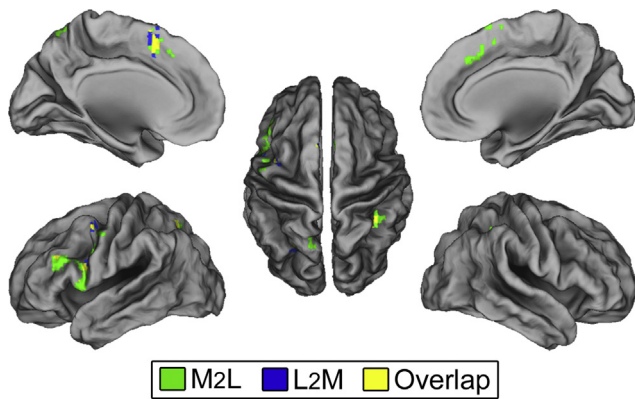
(BA 40) and superior (BA 7) parietal lobuli, as well as in the posterior cerebellum (see Table S2 in the Supplemental Material available online for the complete list of local maxima). As shown in Fig. 2 (regions in yellow), the structure versus repeat contrast uncovered a number of common areas across language and music materials, including the left inferior frontal (in its *pars opercularis*, BA 44) and middle frontal (in BA 6) gyri, as well as the medial frontal/cingulate gyri (BA 6, 32), and bilateral posterior parietal lobe (in both BA 7 and 40). In order to avoid interpreting a “reverse subtraction”, we characterized the mean activity profile for structure and repeat conditions to identify the primary driver in IFG. Mean z-scores from the IFG subregions (defined by external atlases: *pars opercularis* and *pars triangularis* from Harvard-Oxford and *pars orbitalis* from AAL) are displayed in Fig. S2.

The interaction of structure versus repeat structure and materials revealed the left superior and middle temporal gyri (BA 21, 22) to be specific to language (see blue areas in Fig. S1 and Table S3), whereas foci surrounding the right orbital and sub-lobar segments of the inferior frontal gyrus (mainly in BA 47 and 13), along with right superior frontal (BA 6), medial frontal (BA 6, 32) and contralateral posterior cerebellum appeared to be specific to structure in music (see Fig. S1 and Table S3 in the Supplementary Material available online).

#### 3.2. Multivariate analysis

In order to assess whether natural language and music share neural codes for representing structure, we performed a cross-classification multivariate pattern analysis. In this approach, an SVM classifier was trained to recognize structure versus repeat trials in one domain and was then tested on the other. That is to say, we trained a classifier to distinguish structure versus repeat trials in language and then tested it by assessing its ability to discriminate structure versus repeat trials in music (“L2M” cross-classification), and vice versa (“M2L” cross-classification; see Section 2).

As shown in Fig. 3, significant cross-classifications were evident across a number of regions within medial prefrontal cortex, bilateral posterior parietal cortices, as well as left precentral, inferior (in the *pars opercularis*), and middle frontal gyri, matching areas of univariate overlap between the two domains (i.e., yellow regions in Fig. 2). In addition, within each of these cross-classification clusters are areas (in yellow) capable of performing both L2M and M2L classifications, further demonstrating some extent of common underlying neural representation across the two domains. Crucially, this effect could be observed at the single-subject level, with median (single-subject) cross-classification accuracies at 61% for both L2M and M2L classifications, and ranges between 59% and 65%, and 58% and 64% for L2M and M2L classifications, respectively; with chance being 50%). Focusing on the left inferior frontal gyrus in particular, Fig. 4 depicts the reliability of



**Fig. 3.** Group cross-classification (multivariate) result. Green regions represent areas in which the SVM algorithm could significantly classify, with above chance accuracy, structure vs. repeat trials in language materials after having been trained to recognize structure vs. repeat trials in music materials (i.e., M2L cross-classification). Blue regions represent areas in which the SVM algorithm could significantly classify, with above chance accuracy, structure vs. repeat trials in music materials after having been trained to recognize structure vs. repeat trials in language materials (i.e., L2M cross-classifications). Yellow areas show searchlight centers that can significantly perform both classifications.

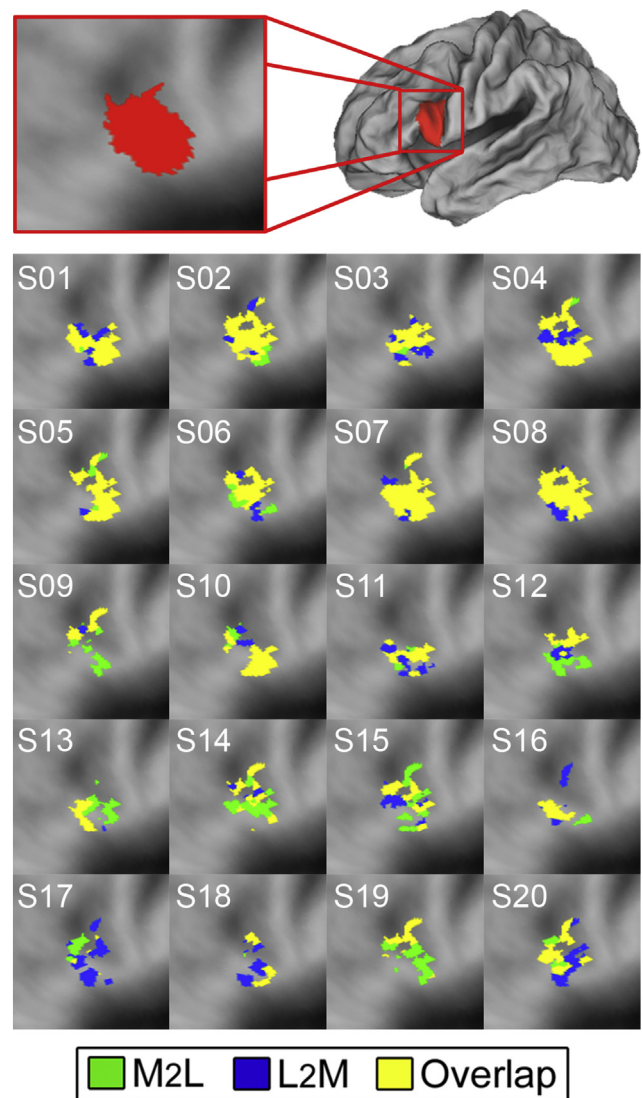
the result at the single-subject level. Fig. 4 also demonstrates a significant across-subject variability in the exact location of voxels sensitive to linguistic structure within the inferior frontal gyrus (as previously shown; Fedorenko, Hsieh, Nieto-Castañón, Whitfield-Gabrieli, & Kanwisher, 2010), something that we also observe in music-structure sensitive voxels, resulting in a systematic but variably located overlap in voxels capable of both L2M and M2L classifications within this region.

We then performed a searchlight classification in order to determine whether Broca’s area was capable of discriminating between the two structured conditions (Active versus Passive voice in Language, and Root vs 2nd Inversion position in Music). After correcting for multiple comparisons, there was no evidence at the group level that regions able to classify the structures within domains were consistent across subjects. However, at the single subject level, a within-subject permutation test revealed that all 20 subjects exhibited, within Broca’s area, regions capable of discriminating between the two language structures ( $L_{AV} \nu L_{PV}$ ; cf. Fig. 5, blue areas), 17 subjects exhibited regions capable of discriminating between the two music structures ( $M_{RP} \nu M_{SI}$ ; cf. Fig. 5, green areas), and for 8 subjects the two sites overlapped (cf. Fig. 5, yellow areas).

**4. Discussion**

In this study we have addressed the question of the relationship between natural language and human cognition by contrasting the neural substrates accompanying the generation of structured sequences across language and music. Overall, our results provide direct evidence for the hypothesis that language and music have a shared neural code for producing structured relationships – a phenomenon that we observe both at the group as well as at the single subject level.

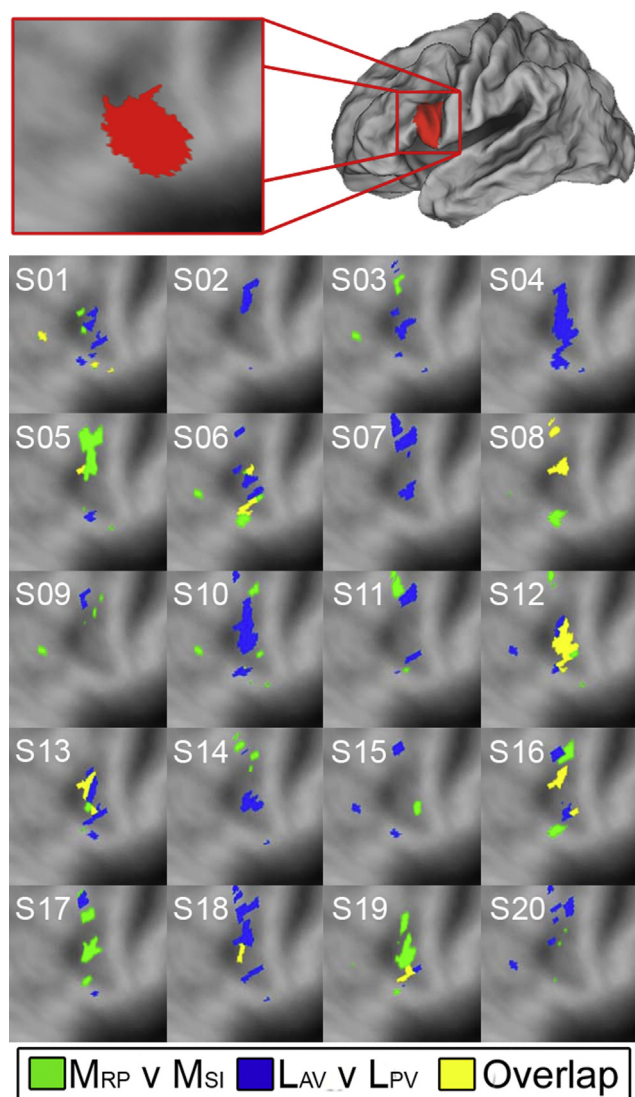
More specifically, we report three central findings. First, employing a magnitude-based univariate approach, we find the generation of structured sequences in language to recruit a well-known left lateralized network of frontal and temporal regions, along with posterior parietal foci, while generation of music sequences engaged a larger, and strongly bilateral, set of fronto-parietal regions. The neural substrate elicited by this performance paradigm (which has remained almost unexplored in the context of music, with the exception of Brown et al., 2006) matches very closely the neural substrate typically reported in tasks focusing on competence in both language (e.g., Ben-Shachar,



**Fig. 4.** Single-subject cross-classification (multivariate) result. Cortical flat-maps depicting, for each participant separately, searchlight centers capable of significant cross-classifications (L2M in blue; M2L in green; overlap in yellow) within the inferior frontal gyrus as defined anatomically (highlighted in red, at the top). Each image (labeled as ‘S##’) represents the classification results for a single participant.

Hendler, Kahn, Ben-Bashat, & Grodzinsky, 2003; Monti, Parsons, & Osherson, 2009) and music (e.g., Maess et al., 2001; Koelsch et al., 2002, 2005; Cheung, Meyer, Friederici, & Koelsch, 2018).

Our second main finding, as evaluated with the same univariate activation-based approach, shows that the building of structured sequences in language and music relies on a number of common regions across left lateral and medial frontal cortices, as well as bilateral posterior parietal regions. In particular, the univariate analysis shows that the posterior-most aspect of Broca’s Area, in the *pars opercularis* of the left inferior frontal gyrus, is metabolically responsive to the presence of structure in the context of both language and music materials (Fig. S2) – a finding that is consistent with results from previous studies (Brown et al., 2006; Koelsch et al., 2002, 2004, 2005; Maess et al., 2001). Beyond the left inferior frontal gyrus, our findings show that the interplay between language and music might well extend to a set of regions in frontal and parietal cortices conventionally referred to as the multiple-demands network (Duncan, 2010), which have been shown to be recruited across a broad class of cognitive operations (Fedorenko, Duncan, & Kanwisher, 2013), and medial frontal regions that are



**Fig. 5.** Single-subject structure classification result. Cortical flat-maps depicting, for each subject, searchlight centers capable of discriminating between structures within each domain ( $L_{AV} v L_{PV}$  in blue,  $M_{RP} v M_{Si}$  in green, Overlap in yellow) within the inferior frontal gyrus.

commonly recruited by motor planning (Haggard, 2008; Tanji and Shima, 1994; Tanji, Shima, & Mushiake, 2007) and that have also been shown to play a role in rhythm perception and generation of spontaneous movement in response to music stimuli (Grahn and Brett, 2007; Lima, Krishnan, & Scott, 2016). The absence of temporal regions identified by the domain-specific main effects (Fig. S1), specifically the involvement of posterior STS (i.e. “Wernicke’s area”) in language tasks but not music, corroborates previous work reporting that posterior temporal regions might engage a semantic/syntactic interaction (see Friederici 2011, 2012 for a review).

Finally, our third, and crucial, finding addresses the significance of the frequently reported overlap between the neural substrate of language and that of music, thereby directly addressing the question of whether the mechanisms of natural language play a role in processing the structured sequences of music. Indeed, while regions of overlapping activation for these two domains have been widely interpreted as marking areas of shared neurocognitive processing (Kunert and Slevc, 2015), these hypotheses had not been directly tested (until now), prompting some to specifically advocate multivariate analyses such as the one adopted here (Peretz et al., 2015). As we reported above, we could find within each of the regions of univariate overlap (in Broca’s

area), areas capable of recognizing music structure on the basis of language structure and vice versa. In fact, in each of these areas a subset of voxels could perform, at the same time, cross-classifications in both directions (i.e.,  $L_{2M}$  and  $M_{2L}$ ), demonstrating a degree of shared neuronal representation of structures across domains. Furthermore, although the result was not significant at the group level, within Broca’s area we could also find, at the single subject level, variably located sub-regions sensitive to the different structures of language (i.e., active versus passive voice) and music (i.e., root versus 2nd inversion position). These findings thus provide evidence in favor of the idea that language and music cognition share, within Broca’s area, resources related to establishing structured relationships tying discrete elements into well-formed complex structures (Patel, 2003, 2012). Nonetheless, our data cannot address the issue of whether the neural representation of specific operations (e.g., syntactic movement, Chomsky, 1957, 1995) can be directly mapped across the two domains, as entailed by some views (Katz & Pesetsky, 2011).

In interpreting our results, however, two important limitations should be noted. First, it is not possible, from our analysis, to determine the precise factor(s) driving the successful classification of either structured versus non-structured trials across domains, or within the structured sequences of each domain (cf., Hebart and Baker 2017). Indeed, in both analyses, stimuli differed across a number of important features including the presence of structured sequences as well as working-memory demands (as shown by the common recruitment of regions within the so-called multiple demands (Duncan, 2010), or extrinsic mode (Hugdahl, Raichle, Mitra, & Specht, 2015), network). The degree to which each of these (and possibly other) factors drives our results cannot be assessed with our own data. Nonetheless, while the co-localization of cross-classifying voxels within Broca’s area is open to such interpretational ambiguity, it is harder to make the same case for the little co-localization observed in the classification of different ordered structures (within each domain). If indeed general processes were responsible for the activations previously observed for music in Broca’s area (Fedorenko & Varley, 2016), a greater co-localization of voxels capable of classifying the two structures within each domain would be expected. Furthermore, a recent study has shown that increased (general) working memory demands are not sufficient to drive activation within Broca’s area, while activating a well-characterized network of frontal and parietal regions (Coetzee & Monti, 2018).

A second important consideration is the fact that, for both music and language, our structures were relatively simple as compared to previous work attempting to relate linguistic and music sequences (e.g., Kunert et al., 2015). It is thus not clear whether the fact that we fail to uncover a group effect in classifying, within each domain, the two structures (i.e., active/passive voice and root/2nd inversion position, for language and music respectively) is due to this difference or other factors such as the high variability of the fine distribution of neural representations across individuals (cf., Fig. 5; see Fedorenko et al., 2010).

Overall, our findings are consistent with the idea that Broca’s area might play a role across multiple domains, in the context of processing structured sequences, as envisioned in the Supramodal Hierarchical Parser hypothesis (cf., Tettamanti & Weniger, 2006). Nonetheless, it is important to note that to date this hypothesis has only found support in a narrow sense (e.g., as conceived in Fadiga et al., 2009, and see Van de Cavey & Hartsuiker, 2016 for evidence of domain-general mechanisms), as it does not appear to extend to the hierarchical relationships of algebra (e.g., Varley et al., 2005; Monti et al., 2012), logic inference (e.g., Monti, Osherson, Martinez, & Parsons, 2007; Monti et al., 2009; Monti & Osherson, 2012), and spatial cognition (e.g., Bek, Blades, Siegal, & Varley, 2010). Thus far, it has only been found to be consistent with findings in the domain of language (cf., Bookheimer, 2002), music (here and in most previous neuroimaging reports; e.g., Maess et al., 2001; Koelsch et al., 2002; Koelsch, Rohrmeier, Torrecuso, & Jentschke, 2013), and motor action sequences (e.g., Fazio et al., 2009). In this

sense, Broca's area cannot be viewed as a central parser capable of operating in any domain of cognition as one would expect of a domain general processor or working memory component, though it may be a core component in a network of regions engaging in hierarchical processing (Fitch, 2014). While it might be speculated that Broca's area plays a role in cognitive domains where structured relationships trigger automatic and effortless intuitions (compare the ease of detecting a non-grammatical sentence or a sour note with the much more taxing task of detecting an incorrect algebraic expression or logic argument), the crucial factor(s) determining its involvement in processing structured sequences remains to be fully specified.

In conclusion, this report provides direct evidence that forging the structured sequences of natural language and music relies also on a common neural representational space which includes Broca's area, a region traditionally associated with the syntactic operations of language. Contrary to recent discussion (Fedorenko & Varley, 2016), we find that activations within Broca's area in response to music stimuli cannot be discounted as epiphenomenal to salient events (e.g., violations, which were not present in our design) or due to recruitment of domain general processes (encapsulated in, or close to, Broca's area). Whether the common mechanisms originally evolved in connection with one of the two domains, or whether they jointly inherited their properties from a common antecedent cognitive domain (e.g., a "protosodic protolanguage," Fitch, 2005; or the capacity for hierarchical sequences of motor actions, Lashley, 1951), remains to be understood.

## 5. Author statement

While Broca's area is a well-established sequence processor in language, whether the region directly computes sequences in other domains is contested. This fMRI study uses production paradigm and cross-classification to show that, at least in the context of language and music, Broca's area represents structured sequences across domains.

## Acknowledgments

This research was supported by the UCLA OVCR-COR Transdisciplinary Seed Grant "Language, Music, and the Brain" to A.L. and M.M.M., the National Defense Science and Engineering (NDSeg) fellowship to J.N.C., and by the Staglin IMHRO Center for Cognitive Neuroscience at UCLA.

## Author contribution

MMM and AL developed the study concept and secured the funding. MMM, MHR, CAB, and DS devised the study design. MHR conducted behavioral testing and neuroimaging data collection. JNC and MHR performed the data analysis and, together with MMM, interpreted the results. MMM drafted the manuscript, MHR and JNC provided critical revisions. All authors contributed to subsequent editing of the manuscript.

## References

Ayotte, J., Peretz, I., & Hyde, K. (2002). Congenital amusia: A group study of adults afflicted with a music-specific disorder. *Brain*, *125*(Pt 2), 238–251.

Ayotte, J., Peretz, I., Rousseau, I., Bard, C., & Bojanowski, M. (2000). Patterns of music agnosia associated with middle cerebral artery infarcts. *Brain*, *123*(Pt 9), 1926–1938.

Bek, J., Blades, M., Siegal, M., & Varley, R. (2010). Language and spatial reorientation: Evidence from severe aphasia. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(3), 646–658. <https://doi.org/10.1037/a0018281>.

Ben-Shachar, M., Hendler, T., Kahn, I., Ben-Bashat, D., & Grodzinsky, Y. (2003). The neural reality of syntactic transformations: Evidence from functional magnetic resonance imaging. *Psychological Science*, *14*(5), 433–440.

Berwick, R. C., Friederici, A. D., Chomsky, N., & Bolhuis, J. J. (2013). Evolution, brain, and the nature of language. *Trends in Cognitive Science*, *17*, 89–98.

Boeckx, C. (2010). *Language in cognition: Uncovering mental structures and the rules behind them*. John Wiley & Sons.

Bookheimer, S. (2002). Functional MRI of language: New approaches to understanding

the cortical organization of semantic processing. *Annual Review of Neuroscience*, *25*, 151–188. <https://doi.org/10.1146/annurev.neuro.25.112701.142946>.

Brown, S., Martinez, M. J., & Parsons, L. M. (2006). Music and language side by side in the brain: A PET study of the generation of melodies and sentences. *European Journal of Neuroscience*, *23*(10), 2791–2803. <https://doi.org/10.1111/j.1460-9568.2006.04785.x>.

Van de Cavey, J., & Hartsuiker, R. (2016). Is there a domain-general cognitive structuring system? Evidence from structural priming across music, math, action descriptions, and language. *Cognition*, *146*, 172–184.

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

Chomsky, N. (1957). *Syntactic structures*. The Hague: Mouton.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, Massachusetts: MIT Press.

Chomsky, N. (1995). *The minimalist program*. Cambridge, MA: MIT Press.

Coetzee, J., & Monti, M. (2018). At the core of reasoning: Dissociating deductive and non-deductive load. *Human Brain Mapping*, *39*(4), 1850–1861. <https://doi.org/10.1002/hbm.23979>.

Ding, N., Melloni, L., Zhang, H., Tian, X., & Poeppel, D. (2015). Cortical tracking of hierarchical linguistic structures in connected speech. *Nature neuroscience*.

Duncan, J. (2010). The multiple-demand (MD) system of the primate brain: Mental programs for intelligent behaviour. *Trends in Cognitive Sciences*, *14*(4), 172–179. <https://doi.org/10.1016/j.tics.2010.01.004>.

Etzel, J., & Braver, T. (2013). MVPA permutation schemes: Permutation testing in the land of cross-validation. *International workshop on pattern recognition in neuroimaging (PRNI)* (pp. 140–143). <https://doi.org/10.1109/PRNI.2013.44>.

Fadiga, L., Craighero, L., & D'Ausilio, A. (2009). Broca's area in language, action, and music. *Annals of the New York Academy of Sciences*, *1169*, 448–458. <https://doi.org/10.1111/j.1749-6632.2009.04582.x>.

Fazio, P., Cantagallo, A., Craighero, L., D'Ausilio, A., Roy, A. C., Pozzo, T., ... Fadiga, L. (2009). Encoding of human action in Broca's area. *Brain*, *132*(Pt 7), 1980–1988. <https://doi.org/10.1093/brain/awp118>.

Fedorenko, E., Duncan, J., & Kanwisher, N. (2013). Broad domain generality in focal regions of frontal and parietal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(41), 16616–16621. <https://doi.org/10.1073/pnas.1315235110>.

Fedorenko, E., Hsieh, P.-J., Nieto-Castañón, A., Whitfield-Gabrieli, S., & Kanwisher, N. (2010). New method for fMRI investigations of language: Defining ROIs functionally in individual subjects. *Journal of Neurophysiology*, *104*(2), 1177–1194. <https://doi.org/10.1152/jn.00032.2010>.

Fedorenko, E., & Varley, R. (2016). The relationship between language and other high-level cognitive functions: Review of neuroimaging and patient evidence. *Annals of the New York Academy of Sciences*.

Fitch, W. T. (2005). The evolution of language: A comparative review. *Biology and Philosophy*, *20*(2–3), 193–203.

Fitch, W. T. (2014). Toward a computational framework for cognitive biology: Unifying approaches from cognitive neuroscience and comparative cognition. *Physics of Life Reviews*, *11*, 329–364.

Fitch, W. T., & Martins, M. D. (2014). Hierarchical processing in music, language, and action: Lashley revisited. *Annals of the New York Academy of Sciences*, *1316*, 87–104. <https://doi.org/10.1111/nyas.12406>.

Friederici, A. D. (2011). The brain basis of language processing: From structure to function. *Physiological Reviews*, *91*(4), 1357–1392. <https://doi.org/10.1152/physrev.00006.2011>.

Friederici, A. D. (2012). The cortical language circuit: From auditory perception to sentence comprehension. *Trends in Cognitive Sciences*, *16*, 262–268.

Gleitman, L., & Papafragou, A. (2013). Relations between language and thought. In D. Reisberg (Ed.), *The Oxford handbook of cognitive psychology* (pp. 504–523). USA: Oxford University Press.

Grahn, J. A., & Brett, M. (2007). Rhythm and beat perception in motor areas of the brain. *Journal of Cognitive Neuroscience*, *19*, 893–906.

Haggard, P. (2008). Human volition: Towards a neuroscience of will. *Nature Reviews Neuroscience*, *9*, 934–946.

Hebart, M. N., & Baker, C. I. (2017). Deconstructing multivariate decoding for the study of brain function. *NeuroImage*. <https://doi.org/10.1016/j.neuroimage.2017.08.005>.

Hugdahl, K., Raichle, M. E., Mitra, A., & Specht, K. (2015). On the existence of a generalized non-specific task-dependent network. *Frontiers in Human Neuroscience*, *9*, 430. <https://doi.org/10.3389/fnhum.2015.00430>.

Jackendoff, R. (2002). *Foundations of language*. New York: Oxford University Press.

Janata, P. (1995). ERP measures assay the degree of expectancy violation of harmonic contexts in music. *Journal of Cognitive Neuroscience*, *7*(2), 153–164. <https://doi.org/10.1162/jocn.1995.7.2.153>.

Katz, J., & Pesetsky, D. (2011). The identity thesis for language and music. Draft published online: [lingbuzz/000959](http://lingbuzz/000959).

Koelsch, S., Fritz, T., Schulze, K., Alsop, D., & Schlaug, G. (2005). Adults and children processing music: An fMRI study. *NeuroImage*, *25*(4), 1068–1076. <https://doi.org/10.1016/j.neuroimage.2004.12.050>.

Koelsch, S., Gunter, T. C., Cramon, D. Y., Zysset, S., Lohmann, G., & Friederici, A. D. (2002). Bach speaks: A cortical "language-network" serves the processing of music. *NeuroImage*, *17*(2), 956–966.

Koelsch, S., Kasper, E., Sammler, D., Schulze, K., Gunter, T., & Friederici, A. D. (2004). Music, language and meaning: Brain signatures of semantic processing. *Nature Neuroscience*, *7*(3), 302–307. <https://doi.org/10.1038/nn1197>.

Koelsch, S., Rohrmeier, M., Torrecuso, R., & Jentschke, S. (2013). Processing of hierarchical syntactic structure in music. *Proceedings of the National Academy of Sciences*, *110*(38), 15443–15448.

Kriegeskorte, N., Goebel, R., & Bandettini, P. (2006). Information-based functional brain

- mapping. *Proceedings of the National Academy of Sciences of the United States of America*, 103, 3863–3868.
- Kunert, R., & Slevc, L. R. (2015). A commentary on: “Neural overlap in processing music and speech”. *Frontiers in Human Neuroscience*, 9, 330. <https://doi.org/10.3389/fnhum.2015.00330>.
- Kunert, R., Willems, R. M., Casasanto, D., Patel, A. D., & Hagoort, P. (2015). Music and language syntax interact in Broca’s area: An fMRI study. *PLoS One*, 10(11), e0141069.
- Lashley, K. (1951). The problem of serial order in behavior. In L. A. Jeffreys (Ed.), *Cerebral mechanisms in behavior; The Hixon symposium* (pp. 112–146). New York: Wiley.
- Lerdahl (2001). *Tonal pitch space*. Oxford University Press.
- Lerdahl, F., & Jackendoff, R. S. (1985). *A generative theory of tonal music*. Cambridge, Mass: MIT Press.
- Luria, A. R., Tsvetkova, L. S., & Futer, D. (1965). Aphasia in a composer. *Journal of the Neurological Sciences*, 2(3), 288–292.
- Lima, C. F., Krishnan, S., & Scott, S. K. (2016). Roles of supplementary motor areas in auditory processing and auditory imagery. *Trends in Neurosciences*, 39, 527–542.
- Maess, B., Koelsch, S., Gunter, T. C., & Friederici, A. D. (2001). Musical syntax is processed in Broca’s area: An MEG study. *Nature Neuroscience*, 4(5), 540–545. <https://doi.org/10.1038/87502>.
- Marin, O. S., & Perry, D. W. (1999). Neurological aspects of music perception and performance. In D. Deutsch (Ed.), *The psychology of music* (pp. 653–724). (2nd ed.). Academic Press.
- Maruyama, M., Pallier, C., Jobert, A., Sigman, M., & Dehaene, S. (2012). The cortical representation of simple mathematical expressions. *Neuroimage*, 61(4), 1444–1460. <https://doi.org/10.1016/j.neuroimage.2012.04.020>.
- Monti, M. M. (2011). Statistical analysis of fMRI time-series: A critical review of the GLM approach. *Frontiers in Human Neuroscience*, 5, 28. <https://doi.org/10.3389/fnhum.2011.00028>.
- Monti, M. M. (2017). The role of language in structure-dependent cognition. In M. Moody (Ed.), *Neural mechanisms of language*. New York, NY: Springer.
- Monti, M. M., & Osherson, D. N. (2012). Logic, language and the brain. *Brain Research*, 1428, 33–42. <https://doi.org/10.1016/j.brainres.2011.05.061>.
- Monti, M. M., Osherson, D. N., Martinez, M. J., & Parsons, L. M. (2007). Functional neuroanatomy of deductive inference: A language-independent distributed network. *Neuroimage*, 37(3), 1005–1016. <https://doi.org/10.1016/j.neuroimage.2007.04.069>.
- Monti, M. M., Parsons, L. M., & Osherson, D. N. (2009). The boundaries of language and thought in deductive inference. *Proceedings of the National Academy of Sciences of the United States of America*, 106(30), 12554–12559. <https://doi.org/10.1073/pnas.0902422106>.
- Monti, M. M., Parsons, L. M., & Osherson, D. N. (2012). Thought beyond language: Neural dissociation of algebra and natural language. *Psychological Science*, 23(8), 914–922. <https://doi.org/10.1177/0956797612437427>.
- Mumford, J. A., Davis, T., & Poldrack, R. A. (2014). The impact of study design on pattern estimation for single-trial multivariate pattern analysis. *Neuroimage*, 103, 130–138. <https://doi.org/10.1016/j.neuroimage.2014.09.026>.
- Mumford, J. A., Turner, B. O., Ashby, F. G., & Poldrack, R. A. (2012). Deconvolving BOLD activation in event-related designs for multivoxel pattern classification analyses. *Neuroimage*, 59(3), 2636–2643. <https://doi.org/10.1016/j.neuroimage.2011.08.076>.
- Musso, M., Weiller, C., Horn, A., Glauche, V., Umarova, R., Hennig, J., Schneider, A., & Rijntjes, M. (2015). A single dual-stream framework for syntactic computations in music and language. *NeuroImage*, 117, 267–283. <https://doi.org/10.1016/j.neuroimage.2015.05.020>.
- Osherson, D. (1975). Logic and models of logical thinking. *Reasoning: Representation and Process in Children and Adults*, 81–91.
- Patel, A. D. (2003). Language, music, syntax and the brain. *Nature Neuroscience*, 6(7), 674–681. <https://doi.org/10.1038/nn1082>.
- Patel, A. D. (2012). Language, music, and the brain: A resource-sharing framework. In P. Rebuschat, M. Rohrmeier, J. Hawkins, & I. Cross (Eds.), *Language and music as cognitive systems* (pp. 204–223). Oxford, UK: Oxford University Press.
- 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(6), 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), 776–789.
- Peirce, J. W. (2008). Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics*, 2, 10. <https://doi.org/10.3389/neuro.11.010.2008>.
- Peretz, I. (1993). Auditory atonalia for melodies. *Cognitive Neuropsychology*, 10(1), 21–56.
- Peretz, I., & Coltheart, M. (2003). Modularity of music processing. *Nature Neuroscience*, 6(7), 688–691.
- Peretz, I., Kolinsky, R., Tramo, M., Labrecque, R., Hublet, C., Demeurisse, G., & Belleville, S. (1994). Functional dissociations following bilateral lesions of auditory cortex. *Brain*, 117(Pt 6), 1283–1301.
- Peretz, I., Vuvan, D., Lagrois, M.-É., & Armony, J. L. (2015). Neural overlap in processing music and speech. *Philosophical Transactions of the Royal Society of London. Series B, Biological sciences*, 370(1664), 20140090. <https://doi.org/10.1098/rstb.2014.0090>.
- Piccirilli, M., Sciarra, T., & Luzzi, S. (2000). Modularity of music: Evidence from a case of pure amusia. *Journal of Neurology, Neurosurgery and Psychiatry*, 69(4), 541–545.
- Rissman, J., Gazzaley, A., & D’Esposito, M. (2004). Measuring functional connectivity during distinct stages of a cognitive task. *Neuroimage*, 23(2), 752–763.
- Rogalsky, C., Rong, F., Saberi, K., & Hickok, G. (2011). Functional anatomy of language and music perception: Temporal and structural factors investigated using functional magnetic resonance imaging. *Journal of Neuroscience*, 31(10), 3843–3852. <https://doi.org/10.1523/JNEUROSCI.4515-10.2011>.
- Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E., Johansen-Berg, H., ... Matthews, P. M. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage*, 23, S208–S219.
- Steinbeis, N., & Koelsch, S. (2008). Shared neural resources between music and language indicate semantic processing of musical tension-resolution patterns. *Cerebral Cortex*, 18(5), 1169–1178. <https://doi.org/10.1093/cercor/bhm149>.
- Tanji, J., & Shima, K. (1994). Role for supplementary motor area cells in planning several movements ahead. *Nature*, 371, 413–416.
- Tanji, J., Shima, K., & Mushiake, H. (2007). Concept-based behavioral planning and the lateral prefrontal cortex. *Trends in Cognitive Sciences*, 11.
- Tettamanti, M., & Weniger, D. (2006). Broca’s area: A supramodal hierarchical processor? *Cortex*, 42(4), 491–494.
- Tillmann, B., Janata, P., & Bharucha, J. J. (2003). Activation of the inferior frontal cortex in musical priming. *Annals of the New York Academy of Sciences*, 999, 209–211.
- Varley, R. A., Klessinger, N. J. C., Romanowski, C. A. J., & Siegal, M. (2005). Agrammatic but numerate. *Proceedings of the National Academy of Sciences of the United States of America*, 102(9), 3519–3524. <https://doi.org/10.1073/pnas.0407470102>.
- Winkler, A. M., Ridgway, G. R., Webster, M. A., Smith, S. M., & Nichols, T. E. (2014). Permutation inference for the general linear model. *NeuroImage*, 92, 381–397.