Intersubjectivity and Shared Dynamic Structure in Narrative Imaginings to Music [*]

Elizabeth Hellmuth Margulis, Natalie Miller, Nathaniel Mitchell, Mauro Orsini Windholz, Jamal Williams, and J. Devin McAuley

NOTE: The examples for the (text-only) PDF version of this item are available online at:
https://www.mtosmt.org/issues/mto.22.28.4/mto.22.28.4.margulis.php

ABSTRACT: This paper digests a recent body of empirical work on narrative imaginings to music, bringing them into dialogue with existing theoretical frameworks. By examining the stories imagined at three geographic locations using various methods, these studies present converging evidence that intuitive imaginings to music can be strikingly similar across individuals, but only when they share a cultural background. The intersubjectivity of these imaginings extends not just to the stories' content, but also to their dynamic event structure as imagined to unfold across the time course of the music.

[1] Music theory and analysis often presumes an experience in which sound is the focal object of a listener's attention and contemplation. Yet Kassabian observes that most music encountered during daily life in urban, industrialized societies is "listened to without the kind of primary *attention* assumed by most scholarship to date" (2013, xi, emphasis in original). The lack of focal attention doesn't make these experiences unimportant, uninteresting, or impossible to study—quite the opposite. Precisely because they reveal the influence of complex factors—social, personal, psychological, cultural—studying these experiences can illuminate more about broadly relevant questions than accounts that restrict themselves to considered scrutiny of notes, and to individuals invested in such consideration.

[2] Scholars have tended to examine this type of listening most frequently in the context of ambient music (cf. Lanza 2004; Sterne 1997; Szabo 2015), and indeed, this genre's commitment to the idea that music should help people deploy attentional resources elsewhere embodies the paradox at the heart of this kind of listening experience. It's like pacing to try to solve a problem—the solution doesn't emerge from concentrated attention on the footsteps, but rather from the footsteps helping to establish a mental state in which a person can think so deeply about a problem that they forget they're walking. This sort of attentional shift from music to something else occurs not just in the

context of ambient music, but also in response to a wide range of musics across many modes of listening (cf. Tuuri and Eerola 2012). For example, a person might intentionally listen to music to help themselves lift heavier weights or run faster (Terry et al. 2020), or a person might listen to a song and unexpectedly find themselves flooded with cascades of autobiographical memories (Jakubowski and Ghosh 2021). Sloboda 2005 notes that in a large UK survey, people reported using music to cue reminiscence as the most common motivation for listening. As Herbert observes, "such reminiscence would involve an inward focus accompanied by visual imagery" (2012, 14).

[3] Just as music can direct attention to aspects of the perceptual environment, the case of reminiscence makes clear that it can also direct attention to imagined scenarios. But memories are not the only internal imaginings that music can cue. Music listening can also evoke internal experiences ranging from visual imagery to kinesthetic sensations (Küssner and Eerola 2019; Margulis and McAuley 2022; Presicce and Bailes 2019). Accounts of acousmatic listening—wherein sounds are encountered via recording technology that does not capture their original multimodal context—emphasize the capacity of recorded music to engender a focus on the sounds themselves, independent of their actual source (Kane 2014); however, it also allows listeners to supply their own imagined multimodal content, which could include a plausible source (e.g., picturing a bow moving furiously back and forth on stage across a violin's strings) or something more fanciful (such as, for the same passage, a swelling storm or an invading army).

[4] Such imaginings have been identified as one of the key mechanisms for emotional response to music (Juslin and Västfjäll 2008), so why have they received comparatively little attention? For one thing, they depend heavily on culture and previous experience, which might seem to introduce a degree of variability that makes research challenging. But the impact of these factors is not necessarily to generate infinite, irreconcilable, idiosyncratic subjective imaginings from individuals, because the ubiquitous presence of music and multimedia ensures that "through this listening and these responses . . . a nonindividual, not simply human, distributed subjectivity takes place across a network of music media" (Kassabian 2013, xi). Kassabian conceives of this distributed subjectivity as "a field, across which information flows . . . leading to affective responses" (ibid.). Although Kassabian's use of the term distributed subjectivity has many layers, most relevant to the work reported here is the notion that aspects of musical experience extend beyond the individual listening subject. For Kassabian, this motivates a form of music analysis that doesn't return to "discrete individuals." The research reported here works from the opposite side toward a similar conclusion—by examining similarities in the patterns of responses made by individual listeners, it reveals the existence of interpretive currents that flow through experiences that might have felt idiosyncratic to the people having them. But the pervasiveness of mass media influences the musical imagination such that scenarios imagined while listening may not be as idiosyncratic or arbitrary as they can seem. To understand the imaginative listening experiences sustained by listeners without special formal training, music theorists may benefit from looking to models in media studies, and from focusing on the intersubjective commonalities that emerge in the descriptive experiential accounts these listeners provide.

Studying Narratives Imagined During Music Listening

[5] This paper digests a recent body of empirical work on narrative imaginings to excerpts of Chinese and Western instrumental art music (carried out by some of the authors on the current paper—E.H.M., J.W. and J.D.M.—in collaboration with others), tracing how the distributed subjectivity articulated by Kassabian refracts through individuals, and presenting the findings and their implications to the music theory community. The set of studies examined here were not the first to consider people's free response imaginings to music. Tagg and Clarida's seminal study of imaginative responses to music involved playing "title tunes"—music that accompanied the opening credits in various TV shows—to students in large lecture classes and asking them to report on the film or TV scene they imagined accompanying it. In a book-length exploration of responses, they argue that "vernacular musical competence (involving the ability to distinguish instantaneously between, say, a Hollywood and an Italian Western, or between horror and mystery)" has been critically neglected by music scholarship, related perhaps to the disciplinary distance (or perhaps even rift) between musicology and media and communication studies (2003,

10). By examining the commonalities that emerge in a highly unconstrained free response task, Tagg and Clarida lay an essential foundation for more of this kind of work.

[6] Despite the trailblazing nature of Tagg and Clarida's study, methodological problems limit the interpretability of the findings and necessitate continued research. Their surveys were conducted informally within classroom settings, without tracking participant demographics or standardizing instructions. How the task was described to participants varied from iteration to iteration, making it impossible to judge whether response differences stemmed from demographic factors, variations in the way the task was presented, or characteristics of the musical excerpts. Moreover, the studies relied on title tunes associated with specific media contexts, and although the researchers tried to select ones that would be unfamiliar to participants, the possibility that participants had previously been exposed to the excerpts within the context of their associated television programs cannot be ruled out.

[7] Most seriously, the survey was administered in classes taught by the experimenter, introducing two key confounds. First, demand characteristics (Nichols and Maner 2008), whereby participants deduce the goal of the experiment and, without even knowing they're doing it, attempt—compliantly—to provide the expected results. This risk is particularly high in the context of a class taught by the experimenter where related discussions might have led into the session, because in this case the experimenter has additional authority over the students as their professor and because the experimenter might have talked about related topics previously with the students, making their intent with the study even easier to guess.

[8] Second, the point of the study is to examine interrelationships that might arise among individual responses, but these interrelationships are only interesting insofar as they have their source in responses that were independently generated. For example, if one student described the story they imagined out loud, and features from their description subsequently appeared in other students' responses, this would not be a particularly interesting demonstration of intersubjectivity. Any recurrence of elements from that story in the other students' descriptions could be explained by the influence of the first student's report. In large lecture classes where students are seated near one another in the same space, it's likely that wandering eyes and the influence of what neighboring students were writing down affected the responses. Together, these potential confounds render the authors' analyses less persuasive.

[9] **Table 1** provides an overview of a more recent body of work that uses an updated set of methods to explore the stories listeners imagine in response to instrumental art music. In this set of studies, participants take part in individual sessions conducted in a sound attenuated booth with no experimenter present, or—in communities that rely on spoken rather than written language—in sessions with a research assistant and translator where the music was presented individually over headphones such that neither the assistant nor the translator knew the identity of the excerpt to which the participant was responding.

[10] Furthermore, in these studies, participant demographics were not only carefully recorded, but also explicitly varied, in the sense that experiments were conducted at multiple geographic locations with the express purpose of understanding the role of cultural experience and exposure in shaping imaginative response. Experiments primarily took place at two land-grant universities in the US Midwest, the University of Arkansas and Michigan State University, where participants had enormous amounts of exposure to Western media, and at a Dong village in a rural part of China—Dimen—where participants had little such exposure. The rationale behind this distribution of study locations was to try to begin to understand the role of media exposure in musical imaginings by, on the one hand, including two geographically distinct groups of participants who, despite their geographic distance, shared broad patterns of media exposure, and, on the other, including two geographically distinct groups whose patterns of media exposure differed.[1]

[11] In addition, this new set of studies used multiple methods not employed in previous work to provide converging evidence about narrative imaginings to music. Tools from machine learning were used to analyze the free response accounts of imagined stories. A real-time task in which participants clicked to indicate *when* they imagined new narrative events in the ongoing imagined

story as the music progressed allowed for inspection of the dynamic time course of narrative imaginings. Participants also completed other tasks for the same set of experiments, such as continuous ratings of perceived tension, permitting a richer understanding of how narrative imaginings relate to other, more extensively studied aspects of music perception. A matching task lowered the bar from spontaneous narrative generation to identifying which of two alternatives was the narrative previous participants tended to provide in response to that excerpt, making it possible to uncover common associations that may not have appeared in a task requiring a free response. Finally, experimentally manipulating attributes of the excerpts allowed a deeper understanding of which characteristics convey narrative implications most clearly.

[12] Note that both the Tagg and Clarida studies and the ones described here focus on narrative imaginings sustained in response to recorded music. Music mediates social interactions in a much wider range of ways (cf. DeNora 2000; Turino 2008). Some of the tools used in the studies listed in Table 1 could potentially be adapted to understand the relationship between music and storytelling in the enactive social sphere.

Dimensions of Narrative Response

[13] Scholars disagree about the definition of narrative, with some postulating that it must include a setting, an agent with a goal, and a potential action facilitated or thwarted by circumstances (Barthes 1975; Herman 2007). For the purposes of this project, we adopted Abbot's minimalist definition of narrative: the representation of an event or a series of events (Abbot 2008).

[14] Across the project (summarized in Table 1), it became clear that musical excerpts vary in at least three measurable ways in the kind of narrative response they elicit. First, they vary in *narrativity*—the likelihood that any individual listener will imagine a story at all. Some excerpts reliably elicit imagined stories from a high percentage of listeners; others much less so (Margulis 2017). This characteristic is measured by simply asking participants a yes or no question about whether they imagined a story while listening, and tallying the percentage who responded yes.

[15] Second, they vary in the degree of *narrative engagement* they elicit—the degree to which any imagined story seems vivid and absorbing (McAuley et al. 2021a). This characteristic is measured with a four-question instrument developed and validated by Margulis et al. 2019. When an excerpt elicits high narrative engagement, it means that the imagined story came to mind effortlessly, played out as the music progressed, and was especially vivid, featuring clear characters, events, and settings.

[16] Third, excerpts vary in the *content* of the imagined stories they trigger. For example, one story's content might involve soldiers marching off to battle, another might involve aliens exploring new galaxies, and a third might involve a little girl chasing a cat. This characteristic is measured by examination of the free response descriptions participants provide of the stories they imagine. When tools from machine learning are used to analyze the similarity of these stories (Margulis et al. 2022a), it reveals that some excerpts elicit imagined story content that is broadly shared among participants and others less so.

[17] Importantly, these characteristics are all separable from the others. One excerpt might elicit a broadly similar story (highly shared *content*) for many listeners (high *narrativity*) but that narrative might be quite faint (low *narrative engagement*). Another might elicit an imagined narrative for only a small subset of listeners (low *narrativity*), but the imagined storyline might be very consistent among them (highly shared *content*) and extremely vivid and gripping (high *narrative engagement*). By studying how narrative responses vary along these dimensions to different musical excerpts for participants with different musical experiences, the underlying processes that give rise to narrative imaginings to music can begin to be understood.

[18] The next few sections highlight the project's findings that are most relevant to music theory.

Initial Findings

[19] The first set of responses were collected at all three of the project's geographic locations—Arkansas, Michigan, and Dimen—for a large pool of 128 excerpts, 64 drawn from Western instrumental art music and 64 from Chinese instrumental art music (Margulis et al. 2019). Arkansas and Michigan participants were broadly familiar with Western instrumental music from media exposure but did not typically listen to this kind of music on a voluntary basis. Similarly, Dimen participants were broadly familiar with Chinese instrumental music from media exposure but did not typically listen to it on a voluntary basis. Although it's impossible to choose materials in cross-cultural contexts that function exactly the same way for each group, this roughly comparable level of exposure was intended as a pragmatic approach to selecting the types of musical examples for the study. Because participants at all groups encountered the musical styles used in the study in their everyday lives via ambient exposure rather than voluntary, focused listening, the excerpts were well suited to pick up on associations shaped by widespread cultural currents.

[20] Participants at all three locations narrativized frequently and to similar degrees while listening. They narrativized not only to music with which they were more familiar (Western for the Arkansas and Michigan participants and Chinese for the Dimen participants), but also to music with which they were less familiar; however, they narrativized comparatively more to excerpts from the more familiar musical tradition. In all three groups, and across all excerpt types, narrativization was correlated with enjoyment. The more participants narrativized to a particular excerpt, the more they reported enjoying that excerpt. It could be that enjoying an excerpt led to increased attention, which made narrativization easier, or it could be that the process of narrativization is inherently pleasurable. Regardless, the fact that people readily narrativized to instrumental music and found the experience enjoyable at all geographic locations laid a foundation for further work. Data from this phase of the project were used to norm all excerpts for familiarity and enjoyment across research sites so that these attributes couldn't serve as confounds in subsequent phases of the study.

[21] One other interesting finding emerged from this first set of studies. The excerpts to which the Arkansas participants were likeliest to narrativize tended to be the same ones in response to which the Michigan participants were likeliest to narrativize. Yet despite the fact that participants in Dimen and Arkansas narrativized to similar degrees overall, there was no relationship between the groups in terms of *which* excerpts tended to trigger this response. This was the first intimation that narrativization might be a cross-culturally relevant mode of music listening, but that the specifics around which excerpts elicit such a response might be culturally bounded.

[22] Note that because the study's design involved directly asking participants about their narrative imaginings in response to music, results can only reveal that people narrativize readily and easily under such conditions; results cannot provide an estimate about how frequently they do so in everyday life. Ongoing work using implicit measures as well as broader queries about music-evoked thoughts seeks to understand how common narrative imaginings might be in daily life; nevertheless, the fact that people can readily imagine stories to music, and that these imagined stories show significant excerpt-specific similarity across listeners makes them an important entryway into understanding how people make sense of the music they hear.

Intersubjectivity in the Content of Imagined Narratives

[23] Coding the free response descriptions of imagined narratives for content in each place, it became apparent that individual excerpts seemed to trigger broadly consistent narratives within each location. Moreover, the narratives for each excerpt in Arkansas and Michigan seemed broadly similar, but the narratives diverged sharply between Arkansas and Dimen and between Michigan and Dimen. That is, common, coherent narrative themes seemed to emerge within excerpts for the Dimen participants, but these were not the same themes that emerged for those excerpts in Arkansas or Michigan.

[24] Consider, for example, an excerpt from the first movement of Webern's String Quartet op. 28. Participants in Arkansas and Michigan seemed to focus on the excerpt's atonality, mapping that feature onto scenarios rife with horror, murder, and paranoia.[2]

- I imagined a murderer in a suburban house trying to find the inhabitant of the house, while she was trying to hide from the murderer.
- I imagined someone in a horror movie when they sense that someone is in there house with them. There is a woman who hears something from another room in the house, or sees a shadow in the corner of her eye, and she goes to look through the house for it.

Participants in Dimen, however, seemed to focus not on the absence of tonality, but rather on the short, staccato articulations and the rapid back-and-forths between high and low register, mapping those onto playful scenarios of fun with friends:

- I play with my favorite friends at home and we go into the mountains, very very happily.
- I see kids in roller skates skating around happily.

[25] These sorts of divergences between the stories generated in Michigan and Arkansas on the one hand, and Dimen on the other hand, seem to suggest that the excerpts might cue very different sets of associations in different cultures. But it's a high bar for similarities to emerge on a task as unconstrained as a free response description. A narrower, binary choice task might reveal that the associations reported in each location are *recognizable* as appropriate to participants in the other locations, even if they aren't the ones that came to mind first.

[26] To test this possibility, McAuley et al. 2021b evaluated the ability of participants to correctly identify consensus narratives that emerged from participants in the prior study. Specifically, the narrative descriptions for individual excerpts from the Arkansas participants in the previous study were evaluated to identify the individual response that was most characteristic of the larger pool (i.e., most similar to the other stories provided in response to that excerpt by Arkansas participants). New participants in Arkansas and Dimen then performed a matching task, determining which of two stories was the one previous participants had provided in response to the excerpt. Critically, the foil (incorrect) stories on each trial consisted of the consensus stories provided in response to other excerpts, and these pairings were shuffled for each participant, ensuring that nothing about the stories themselves beyond the plausibility of their match to the excerpt could be driving the results.

[27] Participants in Arkansas could perform this task at near ceiling levels for both Western and Chinese excerpts; it was trivially clear to them which of two stories had been imagined by previous participants in response to individual excerpts, because the previous participants had also come from Arkansas, and shared a broad set of experiences that had likely shaped the associations. Participants in Dimen, however, performed the task at no better than chance levels overall.

[28] One possible explanation was that the format of the matching task did not convey cross-culturally. To rule out this possibility, two identically formatted control matching tasks involving non-narrative descriptors such as loud/soft and fast/slow were administered for the same musical excerpts. The participants performed at near ceiling levels on both of these control tasks, suggesting that the poor performance on the story matching task indicated not a generalized problem with matching tasks, but a legitimate failure of narrative associations that were obvious to within-culture listeners to convey cross-culturally.

[29] A secondary analysis separated the Dimen participants into two groups: those who reported having ever had exposure to Western media, and those who reported never having had exposure to it. Participants with any reported exposure to Western media performed better on the story matching task for the Western excerpts, suggesting that media exposure at least partially drives narrative associations.

[30] A second experiment, also reported in McAuley et al. 2021b, repeated the matching task for participants in Michigan using both Arkansas-generated narratives (as in the first experiment) and Dimen-generated narratives. Participants in Michigan performed just like the participants in Arkansas—identifying with near ceiling-level accuracy which of two stories previous participants had provided for the excerpt—so long as the stories had been generated by participants in

Arkansas. They performed significantly worse on the matching task for Dimen-generated narratives, reinforcing the notion that narrative imaginings reveal strong within-culture intersubjectivity but little similarity across cultures.[(3)]

[31] Just because new participants can successfully identify which of two stories was most plausibly imagined by a previous participant doesn't mean that, given the same music, they would have spontaneously generated a similar story. Forced-choice tasks, such as choosing between two possible narratives, generally make it easier to show a relationship than open-ended tasks, such as freely recounting an imagined narrative. For a more robust test of the notion that within-culture narrative imaginings are characterized by deep intersubjectivity, free response story reports were examined from a total of 622 participants in Arkansas, Michigan, and Dimen. Each participant was tested individually, providing free response accounts of any story they imagined while listening to a set of one-minute excerpts of Western and Chinese instrumental music that had been previously normed for familiarity and enjoyment.

[32] Tools from natural language processing allowed this otherwise intimidatingly large pool of data to be analyzed (for methods details, see Margulis et al. 2022a). All the stories provided by participants at a particular location in response to a particular excerpt were compiled into a narrative document, or *nardoc*. For example, nardoc AC16 contained all the responses from Arkansas (A) participants to Chinese (C) excerpt number 16. Nardoc DW5 contained all the responses from Dimen (D) participants to Western (W) excerpt number 5. Each nardoc was transformed into a feature vector, where each feature corresponded to the term frequency-inverse document frequency (TF-IDF) score of a word, with a high TF-IDF score indicating that a word is uniquely relevant to a particular document. Next, the cosine of the angle between two nardoc feature vectors was measured, with greater cosine similarity indicating more similarity between the corresponding nardocs. **Example 1** reproduces the figure that plots this semantic similarity in two-dimensional space.

[33] Margulis et at. 2022b explores the semantic relationships in more detail. Here, what is particularly germane are the patterns that emerged at the level of the individual excerpt. The collection of stories provided by participants in Michigan to excerpt C16 were so similar to the ones provided by participants in Arkansas that their labels nearly overlap—the same situation arises for excerpt W9. For the clear majority of excerpts, imagined narratives freely described by participants in Arkansas were highly similar to the imagined narratives for those specific excerpts provided by participants in Michigan—this relationship does not hold, however, between Arkansas and Dimen or between Michigan and Dimen.

[34] To assess the significance of the story similarity by excerpt between Arkansas and Michigan participants, we needed a measure of how much similarity would be expected by chance. In other words, how similar would stories be if individual excerpts were not actually driving distinct intersubjectively reliable responses? This null value was calculated in two separate ways, first, from the excerpts themselves. Instead of generating nardocs by compiling stories provided to a single excerpt, excerpt labels were shuffled within each location (Arkansas, Michigan, Dimen) so that the nardocs constituted random grab bags of stories provided to any excerpt. The similarity that arose among stories for individual excerpts could then be compared to the similarity among stories that arose for different excerpts.

[35] A second null value was calculated from a set of control narratives provided in response to a silent period (unprompted by any musical excerpt, participants simply described any story they imagined). The similarity that arose in response to individual excerpts could then be compared to the baseline similarity that existed among stories that had been generated with no musical prompt.

[36] The similarity of narratives imagined to individual excerpts by participants in Arkansas and Michigan, on average, surpassed these thresholds handily, whereas the similarity between narratives imagined in Arkansas and Dimen and Michigan and Dimen did not come close. This constitutes clear evidence, drawn directly from freely provided, open-ended descriptions of imagined narratives, that Kassabian's "distributed subjectivity" (2013) characterizes the musical imaginings of people with shared exposure to mass media. Descriptions that might have been

expected to constitute arbitrary mind-wandering in fact reveal the reliable influence of the music's acoustic features and the shared pool of previous experiences and exposures that participants in Arkansas and Michigan had—without this shared experiential background to draw on, the intersubjective webs fall apart.

The Real-Time Dynamics of Imagined Musical Narrative

[37] Considering these findings, it's possible to argue that the imagined narratives may have been concocted once the excerpts concluded, rather than envisioned more dynamically as the excerpts progressed. To address this possibility, a separate, two-session study run only with participants in Arkansas used a set of real-time tasks. In an initial session, participants did not complete any tasks with "story" or "narrative" in the instructions. Instead, they moved a slider to indicate fluctuating perceptions of tension and relaxation as the music progressed, a task used frequently to register affective response to music in real time (see Farbood 2012 for an overview).

[38] During the second session, participants heard the same set of excerpts again, reporting for each on any story they imagined while listening (the same task used in the studies previously described). But the key innovation of this study was the task participants performed when they listened to the same set of excerpts a final time. On this hearing—while the music progressed— they pressed a button at the moment they imagined a new narrative event to begin. Once the excerpt concluded, they described the event they had imagined at each of these timepoints.

[39] If imagined narratives are atemporal, and do not unfold dynamically as musical excerpts proceed, then the timepoints at which people report imagining story events should be random, distributed evenly throughout the excerpts. Instead, people's event clicks tend to coalesce at discrete individual timepoints for particular excerpts, suggesting that features of the sound shape narrative imaginings in real time as music progresses, and that stories are not merely concocted after the conclusion of the excerpt. What's more, the moments that tended to prompt imagined events for listeners also tended to prompt shifts in their perception of musical tension, as recorded independently during an earlier experimental session before they had been asked to perform any task related to narratives.

[40] **Example 2** tracks average tension ratings across all 60 seconds of an excerpt from the first movement of William Grant Still's Symphony no. 1 as recorded by Neeme Järvi and the Detroit Symphony Orchestra, starting 1:16 into the track. Participants seemed to experience the first half of the excerpt as relatively low in tension, until around 32 seconds, when perceived tension starts to rise sharply before dropping off again around 53 seconds. What's really interesting here is the position of the red dots. These represent timepoints at which event clicks converged—moments when many participants reported a new event starting in their ongoing imagined narrative (cf. Kozak 2015 for an analysis of convergence around the moments listeners gesture in response to music). The larger the dot, the more participants reported the start of an event at that timepoint. The moments at which most participants reported imagining a new event occurred at the beginning of the excerpt, and at the point midway through where the tension suddenly started to increase. The next moment at which new events were most commonly imagined was the point toward the end of the excerpt when the tension started to decrease again. This relationship between perceived changes in musical tension and perceived onsets of imagined narrative events—even when measured in separate sessions weeks apart—suggests that narratives imagined in response to music unfold in real time, with changes in the imagined narrative aligned to changes in the sound.

A Speculative Closer Look

[41] The dynamic succession of perceived events across imagined narratives seems to follow Zbikowski's notion of sonic analogues for dynamic processes (2017). According to Zbikowski, music's patterned sounds often provide analogues to dynamic processes that are extramusical, which can be understood by appropriately encultured listeners and constitute music's grammar and meaning. In response to a one-minute excerpt from the first movement of Liszt's Dante Symphony, as performed by Daniel Barenboim and the Berlin Philharmonic Orchestra, starting

1:13 into the track, participants frequently recounted chase narratives, often—more specifically—a story of a shark chasing an unsuspecting ocean-going person. But more than a general story about a chase, participants tended to recount a specific sequence of chase events, in which the threat emerges, recedes, and then reasserts itself, as shown in **Example 3**.

[42] These structural mappings point to the relevance of musical repetition and divergence in the structuring of imagined narratives. Consider this list of descriptions (presented in their original form with no corrections or edits) participants provided of the stories they imagined while listening to this excerpt:

- I imagined a shark chasing a swimmer. The swimer manages to punch the shark and get away, only to see it again.

- I imagined a shark chasing prey. It begins to chase one prey, and the prey sees the shark and tries to swim away and escape. The shark then eats the prey. The shark then goes after another prey and the same thing hapens.

- I imagined a huge shark approaching a boat and circling it and then hit it a couple times then circling it again

- I imagined a man on the run from the police through the woods. He finds an opening in the woods to stop and rest on some train tracks. After he sits a train comes arund the corner almost killing him, then police dogs start barking closer and closer to him, he then runs onto a bridge to get away, but a car almost hits him.

- I imagined a shark being spotted off the coast of a beach and then it dissappears only later to be seen chasing after someone in the water

- I imagined someone running away from something. Looking around like there were bombs coming from the sky. Th eperson was trying to find a place where he/she could hide. But he/she couldn't. It's like the same scenario was reoucurring couple times in a row, exactly with the same outcome. No place to hide.

- I imagined a cat and mouse. The mouse is almost caught at times but sneaks away from the cats claws

- i imagined the shark from jaws. when the sound rises he is at the surface, and vice versa

- I imagined a chase. Someone is running away from someone or something they are afraid of. There are a couple close calls, but they are never caught.

- I imagined a man in a trench coat sneaking up on a woman. He pulls out a gun and she then realizes that he's behind her. He shoots several times but somehow misses. He then pulls out a knife and chases her.

- I imagined someone being chased by a scarything trying to grab them. When the music escelated, the scary thing grabbed them. They then escaped and repeated the process.

- imagined a sharks stalking a boat and ramming it with is head but the captian makes the shark retreat but he comes back again and he captain has to fight him again

- I imagined a mouse, unknowingly being stalked by a house cat. The cat repeatedly tries to capture the mouse, but everytime he pounces, the mouse just barely evades his claws.

- I imagined a chase scene between a murderer, and someone who isabout to be murdered. In the beginning he man doesn't know he is being persued by a murderer, and the murderer gets the jump on him. The man becomes horrified, and begins trying to escape, which he does. The man is now running away from the murderer, but he gets caught once again.

- I imagined a mouse running away from a bunch of lines. Then he gets surrounded by the lines. He then escapes only to get surrounded again.

- I imagined I was in a dungeon being held captive and every time I would try and escape, someone would come open the door and put me back where I was supposed to be held captive.

- A tornado comes to a town and dimolishes the town then moves on to the next town and dimolishes that town

- I imagine a pirate that has a bunch of people held captive. The captives constantly try to escape but the head pirate is always there to make them go back

These narratives converge not only on the topic of a chase, but also on considerable aspects of its structure—most saliently, that it involves repeated attempts and evasions. As shown in Example 3,

this would seem to correlate with the repetitive structure of the music.

[43] Indeed, an examination of the descriptions provided of individual time-locked events supports this observation. In the event click experiment (Margulis et at. 2022b), participants' descriptions of the event they imagined at the moment when the opening gesture recurred in the music tended to feature the recurrence of an event in the imagined story:

- the cat begins to stalk the mouse once again
- the mouse escapes and beings running again
- they put me back where I originally was
- the cycle repeats
- The person is in the black and white movie and he/she is redoing the hiding twice in a row
- Scary thing chases person again
- The shark searches for more things to eat
- the shark swims away but comes back

Musical recurrence can thus correlate with recurrence within the imagined narratives, further bolstering the supposition that music can shape narrative imaginings dynamically as it proceeds.

[44] As the opposite pole to recurrence, change, departure, and contrast constitute another salient formal dimension that has been theorized to spur narrative imaginings (Almén 2017). These moments of departure, where the music does something new and unexpected in light of its previous content, have been conceptualized by Hatten 1994 to embody within-excerpt markedness, drawing attention by virtue of diverging from what came before. Indeed, returning to the Still excerpt from Example 2, the two-part structure evident from the tension ratings (flat for the first half, steeply curved for the second) mirrors the musical structure of the excerpt, which progresses from a moderately paced passage for clarinet to a sweeping melody in the violins that intensifies and thickens before fading out. Imagined events cluster overwhelmingly around two timepoints: the start of the excerpt, and 33 seconds in, at the entrance of the violins. Descriptions of events perceived at this timepoint in the event click experiment register a climactic surge of narrative energy at this point: a swell of excitement, a burst of flavor, or a physical leap into the air:

- The dancer then meets someone and she falls deeply in love with him happier than she's ever been
- a young prince walks up to her and is the boy of her dreams and asks her to dance
- he begins to dance with a woman on the street
- her fairy godmother shows up
- they go more excited and had more fun
- then they all being to jump largely
- she kisses him and fireworks seem to go off around them
- The circus people come and out and everyone cheers
- Then he began to smell the flowers and was filled with energy and began a brisk walk
- he started to float up
- he inspects the cheese then takes a bite it's delicious
- The kid finds friends and beings to realize that he loves the experience of camp
- guy sees all these good things happening around him and gets even more happy
- the big bird begins to play with the small birds
- woman and man faill im love and spend their life together

[45] A similar responsiveness to novelty and change reveals itself in the event descriptions participants provided on the event-click task for an excerpt from the opening Allegro (2:23 into the track) of Beethoven's Egmont Overture as recorded by Zubin Mehta and the Los Angeles Philharmonic. Participants reported perceived events most frequently at two timepoints. The first was at the beginning or five seconds in—at the moment when the transition from the introduction lands on a sforzando tonic (upbeat to m. 29). For the majority of excerpts, the opening itself typically functions as a commonly reported event onset; but here, the initial event was often delayed by five seconds, demonstrating that listeners were sensitive to the transitional, introductory character of the first few measures. The second most commonly marked timepoint for an event onset was 41 seconds in, at the moment this opening theme returns in a fortissimo orchestral tutti (upbeat to m. 59), the culmination of a prolonged crescendo on the dominant. People tended to imagine stories with plotlines involving increasing menace, with the specific events recounted at the 41-second timepoint registering the most threatening event:

- Then the evil witch burst in to the wedding and ruined the whole thing puting a curse on them and their family.
- The boat tips over and crashes from the storm and the people go overboard
- The police officer notriced a burglar that stearted running and decided to go after him.
- then the wife runs off with her secret lover
- guy steals purse
- the swordsman approaches his foe things get intense
- A turn down a dark hallway as the POV speeds up faster and faster with every turn.
- Climax of the war occurs and the bad guys are winning
- they start setting the buildings on fire
- He starts getting attacked by cannon balls
- precise moment phantom scares her and she runs away
- she pleads but he refuses to let her go
- a massive beast is revealed

These real-time event markings and descriptions suggest a robust, intersubjective pairing between musical features and imagined events in ongoing imagined narratives that supports work from music theory, semiotics, musicology, and music studies on meaning making in music. Hopefully, empirical approaches can provide useful evidence to corroborate and constrain such theories, refracting ideas through multiple disciplinary lenses to arrive at more useful and nuanced notions of how narrative imaginings to music emerge.

Moving Forward

[46] Research using multiple methods across multiple geographic locations suggests that narrative imaginings that can feel quite subjective and idiosyncratic actually tie together listeners with a broadly shared cultural background; provided the same excerpt, they tend to provide highly similar descriptions of their imagined stories. This intersubjectivity can transcend geographic location, but only when the background culture is broadly shared. Intersubjectively stable fields of interpretation form and disintegrate as experiences change. The shared imaginings of listeners in the twenty-first century US, for example, did not tend to match the explicit program or title, when one existed, that had been given to an excerpt centuries before. For example, the section titled Mephistopheles from Liszt's Faust Symphony most frequently elicited imaginings of cartoon cat and mouse chases.

[47] Reception history—in some cases, hundreds of years of use and reuse—shapes listener tendencies in cascades of interpretations and reinterpretations, such that no musico-perceptual experience can exist outside of its sphere of influence. Within this continual flux of use and reuse,

stable points of shared associations can emerge, but these are contingent on time and place. Traditional topic theory, which has focused on the network of associations sustained by listeners in eighteenth-century Europe (cf. Ratner 1980; Mirka 2014), might be conceived as an exploration of one of these stable points of shared associations within the ongoing flux across time and place of music use and reuse. It doesn't matter if Liszt didn't envision the frenzied sparring of a cartoon cat and mouse when penning the Faust Symphony; for a community of people who share pervasive exposure to *Tom and Jerry* (as in the participants from Arkansas and Michigan in the studies reported here), imagery and storylines from these shows can shape the perceptual organization of his music. Fifty years before the same music might have cued an entirely different intersubjective association, and a hundred years before that, yet a different one. Similarly, Baroque orchestral dance music evokes a "fancy," formal setting for many listeners, likely due to influences like the opening of PBS's Masterpiece Theatre and its parodies and imitators. This viewpoint is bolstered by the tendency of listeners at both US research sites to situate these formal settings in the past, but not the eighteenth century specifically— instead, listeners implicate periods they describe variously as "olden times," "the middle ages," "the Victorian era," "the nineteenth century" and "the Renaissance" (Margulis 2017). People's imagined narratives reveal the extent to which musical listening is embedded within rich networks of life experiences, not siphoned off into some abstract and "purely musical" channel, and the extent to which changing life experiences across time, place, and other cultural dimensions can modify these implicit meanings. Music theory that seeks to understand listening should expressly include these factors, rather than assuming a purely musical mode of interaction. Thompson, Bullot, and Margulis (2022) argue that self-oriented responses to music, such as autobiographical memories and identity exploration, as well as source sensitivity, such as a listener's knowledge about the cultural context of music making, are as essential to and definitive of listening experiences as the processing of musical materials and structure. Needless to say, this perspective is well represented in music studies outside of traditional music theory. One of the contributions the studies summarized here make is to reveal that the kind of quantitative approaches music theory has typically privileged can be usefully applied not just to musical structures, but also to interactions between people and musical structures.

[48] Shared narratives seem to emerge from the intersection between a piece's acoustic features and the prior experiences a listener brings to the session—their media exposure, their subject position, their relationship with the apparent source of the music, etc. One approach to understanding the process is to start from the stories people provide and work backwards to understand the influence of both the music's acoustic features and the listener's previous experiences. Another approach is to start by systematically varying the music's acoustic features and examine the effect on the generated stories (McAuley et al. 2021a), or to start by systematically varying the kinds of experiences the participants have previously sustained, and examine the relationship with the stories they produce (Margulis et al. 2019; 2022a). One important methodological commitment of this project is taking seriously the musical experiences of people without specialized formal training. Rather than exploring the way any particular excerpt should be heard, aligning with the prescriptive agenda from some strains of music theory, this enterprise seeks to understand how people do hear it—a decidedly more descriptive undertaking (Temperley 2001).

[49] Because reports of imagined events can map changes as they occur over the course of an excerpt, such imaginings can be understood to have the capacity to track salience. If the acoustic features that are most noticeable to listeners without special training in music shape the structures of the stories they imagine, then these stories can be used to study the features to which they attend, a topic that has been notoriously difficult to examine within the psychology of music.

[50] There seems to be a sweet spot in terms of the amount of change or contrast that will trigger structural changes or event percepts within an imagined narrative. Some kinds of change are too extreme: participants will say, for example, that they'd been imagining a story until the entrance of the throbbing violins, at which point it all disappeared, or a completely new and different story emerged to replace it. Narrative accounts typically proffered in music theory, in contrast, often prioritize unity and continuity, and seek to assimilate diverse elements into the overarching trajectory of a single story. The intuitive narratives people produce suggest instead that listeners are willing to abandon unfolding storylines and pick up new ones as the excerpts proceed.

[51] Since Meyer 1956, musical tension—conceived as arising from the frustration of expectations while listening—has served to transmogrify musical events into felt perceptions. Yet in its typical conception, theory of this sort translates abstract musical phenomena such as pitches and rhythms into similarly abstract psychological phenomena: perceptions of tension and relaxation. The relationship between imagined narrative and tension percepts reveals that these tension responses are in fact couched within larger referential, cinematic, and expressive trajectories that are anything but abstract (see Cox 2016). While canonical accounts of musical tension can seem to occupy a platonic sort of "purely musical" register, the imagined narratives with which experiences of musical tension are clearly linked makes it impossible for music theory to ignore the real-world experiences and imaginings that listeners bring to music. Music theory cannot explain them without studying people with the same attention and ingenuity it has used to understand scores.

Elizabeth Hellmuth Margulis
Princeton University
217 Woolworth
Princeton, NJ 08544
margulis@princeton.edu

Natalie Miller
Princeton University
217 Woolworth
Princeton, NJ 08544
npmiller@princeton.edu

Nathaniel Mitchell
University of North Carolina Greensboro
The College of Visual and Performing Arts
100 McIver St, Greensboro, NC 27412
ndmitchell@uncg.edu

Mauro Orsini Windholz
Princeton University
217 Woolworth
Princeton, NJ 08544
windholz@princeton.edu

Jamal Williams
Princeton University
217 Woolworth
Princeton, NJ 08544
jamalw@princeton.edu

J. Devin McAuley
Michigan State University
College of Natural Science
316 Physics Rd
East Lansing, MI 48824
dmcauley@msu.edu

Works Cited

Abbot, H. Porter. 2008. *The Cambridge Introduction to Narrative*. Cambridge University Press. https://doi.org/10.1017/CBO9780511816932.

Almén, Byron. 2017. *A Theory of Musical Narrative*. Indiana University Press. https://doi.org/10.2307/j.ctt2005szf.

Barthes, Roland. 1975. "An Introduction to the Structural Analysis of Narrative." Translated by Lionel Duisit. *New Literary History* 6 (2): 237–72. https://doi.org/10.2307/468419.

Cox, Arnie. 2016. *Music and Embodied Cognition: Listening, Moving, Feeling, and Thinking*. Indiana University Press. https://doi.org/10.2307/j.ctt200610s.

DeNora, Tia. 2000. *Music in Everyday Life*. Cambridge University Press. https://doi.org/10.1017/CBO9780511489433.

Farbood, Morwaread M. 2012. "A Parametric, Temporal Model of Musical Tension." *Music Perception* 29 (4): 387–428. https://doi.org/10.1525/mp.2012.29.4.387.

Hatten, Robert S. 1994. *Musical Meaning in Beethoven: Markedness, Correlation, and Interpretation*. Indiana University Press.

Herbert, Ruth. 2012. *Everyday Music Listening: Absorption, Dissociation and Trancing*. Routledge.

Herman, David, ed. 2007. *The Cambridge Companion to Narrative*. Cambridge University Press. https://doi.org/10.1017/CCOL0521856965.

Jacoby, Nori, Elizabeth H. Margulis, Martin Clayton, Erin Hannon, Henkjan Honing, John Iversen, Tobias R. Klein, Sam A. Mehr, Lara Pearson, Isabelle Peretz, Marc Perlman, Rainer Polak, Andrea Ravignani, Patrick E. Savage, Gavin Steingo, Catherine J. Stevens, Laurel Trainor, Sandra Trehub, Michael Veal, and Melanie Wald-Fuhrmann. 2020. "Cross-Cultural Work in Music Cognition." *Music Perception* 37 (3): 185–195. https://doi.org/10.1525/mp.2020.37.3.185.

Jakubowski, Kelly, and Anita Ghosh. 2021. "Music-Evoked Autobiographical Memories in Everyday Life." *Psychology of Music* 49 (3): 649–66. https://doi.org/10.1177/0305735619888803.

Juslin, Patrik N., and Daniel Västfjäll. 2008. "Emotional Responses to Music: The Need to Consider Underlying Mechanisms." *Behavioral and Brain Sciences* 31 (5): 559–75. https://doi.org/10.1017/S0140525X08005293.

Kane, Brian. 2014. *Sound Unseen: Acousmatic Sound in Theory and Practice*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199347841.001.0001.

Kassabian, Anahid. 2013. *Ubiquitous Listening: Affect, Attention, and Distributed Subjectivity*. University of California Press. https://doi.org/10.1525/california/9780520275157.001.0001.

Kozak, Mariusz. 2015. "Listeners' Bodies in Music Analysis: Gestures, Motor Intentionality, and Models." *Music Theory Online* 21 (3). https://doi.org/10.30535/mto.21.3.7.

Küssner, Mats B., and Tuomas Eerola. 2019. "The Content and Functions of Vivid and Soothing Visual Imagery During Music: Findings from a Survey Study." *Psychomusicology: Music, Mind, and Brain* 29 (2–3): 90–99. https://doi.org/10.1037/pmu0000238.

Lanza, Joseph. 2004. *Elevator Music: A Surreal History of Muzak, Easy-Listening, and Other Moodsong*. University of Michigan Press. https://doi.org/10.3998/mpub.8718.

Margulis, Elizabeth H. 2017. "An Exploratory Study of Narrative Experiences of Music." *Music Perception*, 35 (2), 235–48. https://doi.org/10.1525/mp.2017.35.2.235.

—————. 2022. "Using Music to Probe How Perception Shapes Imagination." *Trends in Cognitive Sciences*. Advance online publication. https://doi.org/10.1016/j.tics.2022.07.011.

Margulis, Elizabeth H., Patrick C.M. Wong, Rhimmon Simchy-Gross, and J. Devin McAuley. 2019. "What the Music Said: Narrative Listening Across Cultures." *Palgrave Communications* 5 (1): 146. https://doi.org/10.1057/s41599-019-0363-1.

Margulis, Elizabeth H., Jamal Williams, Rhimmon Simchy-Gross, and J. Devin McAuley. 2022a. "When Did That Happen? The Dynamic Unfolding of Perceived Musical Narrative." *Cognition* 226

(September): 105180. https://doi.org/10.1016/j.cognition.2022.105180.

Margulis, Elizabeth H., Patrick C.M. Wong, Cara Turnbull, Benjamin Kubit, and J. Devin McAuley. 2022b. "Narratives Imagined in Response to Instrumental Music Reveal Culture-Bounded Intersubjectivity." *Proceedings of the National Academy of Sciences* 119 (4): e2110406119. https://doi.org/10.1073/pnas.2110406119.

McAuley, J. Devin, Patrick C.M. Wong, Lucas Bellaiche, and Elizabeth H. Margulis. 2021a. "What Drives Narrative Engagement with Music?" *Music Perception* 38 (5): 509–21. https://doi.org/10.1525/mp.2021.38.5.509.

McAuley, J. Devin, Patrick C.M. Wong, Anusha Mamidipaka, Natalie Phillips, and Elizabeth H. Margulis. 2021b. "Do You Hear What I Hear? Perceived Narrative Constitutes a Semantic Dimension for Music." *Cognition* 212 (July): 104712. https://doi.org/10.1016/j.cognition.2021.104712.

Meyer, Leonard B. 1956. *Emotion and Meaning in Music*. The University of Chicago Press.

Mirka, Danuta, ed. 2014. *The Oxford Handbook of Topic Theory*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199841578.001.0001.

Nichols, Austin Lee, and Jon K. Maner. 2008. "The Good-Subject Effect: Investigating Participant Demand Characteristics." *The Journal of General Psychology* 135 (2): 151–66. https://doi.org/10.3200/GENP.135.2.151-166.

Presicce, Graziana, and Freya Bailes. 2019. "Engagement and Visual Imagery in Music Listening: An Exploratory Study." *Psychomusicology: Music, Mind, and Brain* 29 (2–3): 136–55. https://doi.org/10.1037/pmu0000243.

Ratner, Leonard G. 1980. *Classic Music: Expression, Form, and Style*. Schirmer Books.

Savage, Patrick E., Nori Jacoby, Elizabeth H. Margulis, Hideo Daikoku, Manuel Anglada-Tort, Salwa E.-S. Castelo-Branco, Florence E. Nweke, Shinya Fujii, Shantala Hegde, Hu Chuan-Peng, Jason Jabbour, Casey Lew-Williams, Diana Mangalagiu, Rita McNamara, Daniel Müllensiefen, Patricia Opondo, Aniruddh D. Patel, and Huib Schippers. 2021. "Building Sustainable Global Collaborative Networks: Recommendations from Music Studies and the Social Sciences. Preprint. PsyArXiv. https://doi.org/10.31234/osf.io/cb4ys.

Sloboda, John. 2005. *Exploring the Musical Mind: Cognition, Emotion, Ability, Function*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198530121.001.0001.

Sterne, Jonathan. 1997. "Sounds like the Mall of America: Programmed Music and the Architectonics of Commercial Space." *Ethnomusicology* 41 (1): 22–50. https://doi.org/10.2307/852577.

Szabo, Victor. 2015. "Ambient Music as Popular Genre: Historiography, Interpretation, Critique." PhD diss., University of Virginia.

Tagg, Philip, and Bob Clarida. 2003. *Ten Little Title Tunes: Towards a Musicology of the Mass Media*. The Mass Media Music Scholar's Press.

Temperley, David. 2001. "The Question of Purpose in Music Theory: Description, Suggestion and Explanation." *Current Musicology* 66 (Spring): 66–85.

Terry, Peter C., Costas I. Karageorghis, Michelle L. Curran, Olwenn V. Martin, and René L. Parsons-Smith. 2020. "Effects of Music in Exercise and Sport: A Meta-Analytic Review. *Psychological Bulletin* 146 (2): 91–117. https://doi.org/10.1037/bul0000216.

Thompson, William F., Nicolas J. Bullot, and Elizabeth H. Margulis. 2022. "The Psychological Basis of Music Appreciation: Structure, Self and Source." *Psychological Review*. Advance Online Publication. https://doi.org/10.1037/rev0000364.

Turino, Thomas. 2008. *Music as Social Life: The Politics of Participation*. The University of Chicago Press.

Tuuri, Kai, and Tuomas Eerola. 2012. "Formulating a Revised Taxonomy for Modes of Listening."
   Journal of New Music Research 41 (2): 137–52. https://doi.org/10.1080/09298215.2011.614951.

Zbikowski, Lawrence M. 2017. *Foundations of Musical Grammar*. Oxford University Press.
   https://doi.org/10.1093/oso/9780190653637.001.0001.

Footnotes

1. The challenges and potential that emerged as this cross-cultural work developed inspired the organization of two workshops devoted to thinking through how to best carry out such research, culminating in two papers that combine voices from multiple fields to produce some suggestions about best practices (Jacoby et al. 2020, Savage et al. in press).
Return to text

2. Throughout, responses are copied verbatim, including typos and errors.
Return to text

3. A study planned for May of 2020 involved testing story matching for Dimen-generated narratives in Dimen; unfortunately, the pandemic made it impossible to carry out that planned phase of research.
Return to text

Copyright Statement

SOCIETY FOR MUSIC THEORY