Musical Repetition Detection Across Multiple Exposures
Author(s): Elizabeth Hellmuth Margulis
Published by: University of California Press

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at http://www.jstor.org/page/info/about/policies/terms.jsp

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

University of California Press is collaborating with JSTOR to digitize, preserve and extend access to Music Perception: An Interdisciplinary Journal.
Music is a strikingly repetitive stimulus, especially when compared with a domain like language. Ethnomusicologist Bruno Nettl (1983) identifies repetition as a “musical universal” – a characteristic shared by all known musical cultures – and evolutionary biologist W. Tecumseh Fitch calls it a “design feature” of music (2006). As a composer explicitly concerned with generating a sense of structure in music, Arnold Schoenberg conceded that “intelligibility in music seems to be impossible without repetition” (1967). Indeed, as Kivy (1993) notes, musical notation itself is rife with symbols that instruct players to repeat – from the repeat sign to the tremolo to simile marks and da capo.

The perceptibility of musical repetition varies according to the nature of the repeating element and its context. Borges’ Pierre Menard, Author of the Quixote is fiction in the style of a review of a word for word rewriting of chapters from Cervantes’ Don Quixote by the (fictional) contemporary writer Pierre Menard. We are told that “the text of Cervantes and that of Menard are verbally identical, but the second is almost infinitely richer” (p. 52). The same could be said of many reprises in familiar repertoire (the end of Götterdämmerung, with the recurrence of the redemption motive, the end of the Goldberg Variations, with the repeat of the aria), where intervening context, or even perhaps the absence of intervening context, conspires to make something that is ostensibly the same sound very different. Cone puts so much stock in changing context that he believes “in general, there is no such thing as true redundancy in music” (1989, p. 46). Yet Kivy (1993) protests: nevertheless, things repeat. There are important senses in which context reconures sound such that no repetition is truly redundant, but we have to account for the fact that we do experience many reoccurrences as “repetition.” Hanninen (2003) terms this double function recontextualization.

To hear something as a repetition, you must first hear it as a something. In this way, repetition detection can be a useful methodology to investigate perceptual units.
Hanninen (2003) states, "Repetition presumes recognition of a 'thing' that is repeated; to recognize this 'thing,' we must abstract the 'thing' from its context" (p. 59). This study represents a preliminary attempt to investigate the circumstances and contexts within which such "things" are abstracted. What makes a repetition easy to detect and what makes it difficult (keeping in mind that composers may not want all repetitions to be consciously recognized)? How much of this effect is due to the nature of the thing being repeated (a two-note motive or an eight-measure phrase?) and how much of it is due to context (whether the repetition occurs immediately or after intervening material, for example)? Furthermore, how do the effects of these factors change across multiple exposures — when a listener hears a piece not once, but several times?

It is important to note at the outset that this study uses the word "repetition" in a highly constrained way — to designate cases in which an exact sequence of notes and rhythms recurs within an excerpt. These elements could range from short two-note motives to longer phrases and sections; however, no modified repetitions or transformations were considered in this study, in an attempt to keep things as simple as possible during the initial application of a new methodology. This terminological usage is purposefully obtuse and amusical in some ways — the idea is to take a highly literal, algorithmically capturable idea of repetition and use it as a starting point for investigating the more rich and subtle perceptual world of repetition as actually experienced and conceptualized by human listeners.

Consider one of the excerpts used in this study, the fourth piece from Schumann’s Albumblätter, reproduced here as Figure 1. Although it’s possible to think about mm. 9-16 as a parallel consequent to the antecedent mm. 1-8, the literal definition employed here recognizes only that mm. 9-11 repeat mm. 1-3. It is possible to look at the detection rates for this repetition and infer information about people's understanding of musical form, but to collate literal repetition and detection success, only this simple definition was employed. For a more nuanced, but still algorithmically capturable definition of repetition, see Steedman (1977), and for an investigation of the ways the concept of an abstract "theme" emerges from its surface manifestations, see Pollard-Gott (1983) and Deliège (1996). Deliège focuses on musical "cues" — a set of distinctive features that can occupy any domain (rhythmic, metric, timbral, dynamic), and are significant enough to trigger recognition, despite the possibility of many changes in other parameters. For an extensive investigation of this perspective, see Deliège (2001) and the entire special issue of Music Perception in which it is found, as well as Deliège, Melen, Stammers, and Cross (1996).

Lamont and Dibben (2001) provide an excellent overview of a central contrast in the way similarity and categorization are approached by cognitive psychology. According to prototype theory, similarity can be understood as the distance between a particular item and an abstract prototype. But according to theory-based classification, similarity is highly context-dependent and modulated by knowledge and experience. Zbikowski (2002) aligns with the latter perspective, developing a theory of the way conceptual models are fluidly applied in different musical contexts to create rich networks of motivic similarity. Cambouropoulos (2009) also makes a case for the pervasive role of context in perceptions of musical similarity. Deliège’s empirical work on musical similarity examines the way rapidly detected and perceptually salient local features function as "cues" and feed into the segmentation of the musical surface. According to this account, motifs are perceived as similar when they share cues. Empirical work on this theory suggests that at least five factors affect the processes of cue abstraction, causing it to vary from one listener to another: experience/age, familiarity with the musical sample, familiarity with the musical style, complexity of the musical sample, and complexity of the task a listener is asked to perform. Since the present study examines perception across multiple exposures, findings regarding the impact of familiarity with the musical sample are particularly relevant. Bartlett and Dowling (1980) and Dowling and Bartlett (1981) found that surface features such as loudness and texture influence similarity judgments most in contexts where the music is unfamiliar, as in the case of an initial hearing, and less in contexts where the music is more familiar — for example, after several hearings.

The computational problem of how to automatically detect repeating patterns in music is notoriously difficult to solve. Significant attempts have been made by Rolland (1999), Meredith, Lemström, and Wiggins (2002), and Cambouropoulos (2006). Although the present study is not computational in nature, it does seek to identify which repeating patterns listeners recognize as such. It is conceivable that these kinds of empirical data could be incorporated into future computational work that seeks not to automatically extract all cases of repeating patterns, but rather to mimic human behavior and extract the ones that most people seem able to detect (likely a subset of the former, or, in certain interesting cases when listeners perceive repetitions to exist when they in fact do not, an intersecting set).

In much of the empirical literature related to similarity perceptions in music, excerpts were abstracted from the
music and listeners were asked to rate the similarity between them. The present study tackles the topic differently in two ways. First, it addresses perceptions of repetition, rather than similarity. Monaghan and Rowson (2008) tested learning for tone sequences that included different types of patterns, including repetition. Patterns including repetition were learned differently and better than other types that included varying degrees of similarity. Their results support the notion that repetition is a special case for learning, rather than an endpoint on a continuum of similarity. If different perceptual mechanisms indeed underlie the processing of repetition in comparison to the processing of similarity, then it is a particularly important topic to study.

Second, the present study asks listeners to identify repetitions in real time, while the music is ongoing. This adds a degree of ecological validity to the experiment that has not been present in other investigations of the topic. It allows for the examination of processes that underlie the abstraction of segments and the perception of repetition during real musical listening. This issue is critical to an understanding of form as perceived, rather than form as studied or remembered. A better empirical understanding of this issue could feed back into theories of musical form, helping to illuminate their perceptual realities.

Method

Participants
Participants were 29 members of the University of Arkansas community, recruited by flyers hung around campus. There were 13 males and 16 females, and ages ranged from 17 to 36 ($M = 21, SD = 4$). Five participants reported no formal training in music. Participants had a mean of 4.1 years of private lessons ($SD = 4.3$), and 0.9 years of theory training ($SD = 2.7$), but this theory training was concentrated in 5 participants – 24 reported no theory training whatsoever. Participants listened to an average of 21.1 hours of music each week ($SD = 24.2$). They attended an average of 6.3 concerts a year ($SD = 8.4$). The genre that most participants (22/29) reported listening to regularly was rock and alternative music, but 15 of the participants – just over half – reported listening regularly to classical music.

Materials
Testing took place in the music cognition lab at the University of Arkansas. The experiment was programmed in MediaLab (Jarvis, 2006). Stimuli were presented over Sennheiser PXC 450 headphones adjusted to a comfortable listening level, attached to a Dell laptop computer. Stimuli were four short musical excerpts: 1) Rameau, Tambourin in E Minor, performed by Trevor Pinnock, harpsichord; 2) Haydn, Piano Sonata No. 41 in A Major, Hob XVI: 26, II, Menuet al rovescio, performed by Jeno Jando, piano; 3) Schumann, Albumblätter, Op. 124, No. 4, Walzer, performed by Denes Varjon, piano; and 4) Strauss, Der Rosenkavalier, Op. 59, Act II, The Presentation of the Rose, four measures before Rehearsal 25 through Rehearsal 27, performed by Erna Berger, soprano, and the Metropolitan Opera Orchestra, conducted by Fritz Reiner. An additional stimulus was used in the training phase to acclimate participants to the task: Beethoven, String Quartet Op. 18 No. 4, III, mm. 1-50.

Design and Procedure
Participants were seated at a computer terminal. They answered a short questionnaire about their musical background. Next, they were trained on the task and given a practice trial with feedback. They were instructed that they would be asked to identify musical repetitions as they occurred by checking onscreen boxes. They were instructed that anything from a two-note gesture to an entire section, repeated immediately or much later in the excerpt, could count as a repetition, and that they should check the box as quickly as possible once they heard something repeat. They were presented two examples of musical repetitions – one was a short segment that repeated immediately, and one was an entire phrase that repeated after lots of intervening material. Participants were reminded that a repeating element could be short or long, and recur immediately or much later. They then moved on to the experimental trials.

The trials were blocked by excerpt, and these excerpts were presented in random order. Participants heard each excerpt four times within a block, and performed the task (checking boxes every time a repetition occurred) during each of the four exposures. After each exposure, they were asked to enter, in the form of a typed free response, an account of the repetitions they had heard. These responses were used to resolve ambiguities in the analysis of the box checks, and to verify that the boxes were being checked in response to the actual repeating elements in the music. Repetitions identified between 500 and 2000 ms of their onset were taken as correct responses. In ambiguous cases, the free responses were used to identify the repetition to which participants were responding.

Results
Each repeating element in each stimulus excerpt was tallied and encoded. For each of these repeating elements
within each piece, the length of the repeating unit (LRU) was calculated by measuring the time in seconds from the first note of the repeating element to the last. For example, if a four-measure unit repeated, its LRU was the duration in seconds from the first to the last note in that four-measure unit. The mean LRU across all excerpts was 3.07 s (SD = 4.06). The mean LRU for repeating units was 5.44 s (SD = 2.48) in the Strauss excerpt, 4.60 s (SD = 4.73) in the Schumann excerpt, 3.01 s (SD = 4.65) in the Haydn excerpt, and 2.60 s (SD = 2.20) in the Rameau excerpt. These differences were significant $F(3, 6368) = 142.28, p < .001$, with the contrast between LRU s in Rameau and Strauss – the excerpts with the shortest and longest LRU s, respectively – also significant ($p < .001$). LRU is at once a primitive measure, failing to account for musical context or structure, and an illuminating one, since it can expose changes that occur across diverse repertoires and musical situations. Although it is not a perfect encapsulation, it serves here as a reasonable proxy for hierarchic level.

As one indication of the influence of LRU on repetition detection, we tested the hypothesis that exposure improves the probability of a correct response for pieces with a lot of long repeated units (high mean LRU) more than the probability of a correct response for pieces with a lot of short repeated units (low mean LRU). We theorized that repeated exposures might shift attention to increasingly higher levels of the musical structure. Specifically, we compared the probability of a correct response across exposures for repetitions in the Rameau piece, which had relatively short repeated units ($M = 2.60$ s) to repetitions in the Strauss piece, which had relatively long repeated units ($M = 5.44$ s). Figure 2 shows the effect of exposure on the probability of correct responses for repetitions in these pieces.

We tested the model with a repeated measures logistic regression, which controlled for the multiple responses within subjects. Results indicated that more exposures of the piece with the shortest mean LRU (Rameau) led to decreasing odds of a correct response ($OR = 0.92, p = .025$). That is, each additional exposure led to an 8% decrease in the odds of a correct response. In contrast, more exposures in the piece with the longest LRU (Strauss) led to increasing odds of a correct response ($OR = 1.31, p < .001$). That is, each additional exposure led to a 31% increase in the odds of a correct response. The difference in these odds ratios was significant, $p < .001$, as indicated by the interaction.

Since the Rameau and the Strauss differed strikingly in timbre, structure, style, as well as LRU, we used several other methods to investigate the relationship of LRU and exposure, seeking to identify converging evidence about their relationship. We used a repeated measures logistic regression, which controlled for the multiple responses within subjects, to examine the effect of LRU on repetition detection across all excerpts. There was a significant interaction between exposure and LRU ($p < .001$). As Figure 3 illustrates, when LRU was small, exposure reduced the probability of a correct response. When LRU was large, exposure increased the probability of a correct response.

To investigate the effect of LRU on repetition detection across all exposures, we used a repeated measures logistic regression with a quadratic term. Figure 4 shows the model employed in this test – repetition detection was best for LRUs of medium length, around 6 s, and decreased as a LRUs got very short or very long. This trend was significant, $p < .001$.

Some repetitions were immediate – once the repeating unit ended, it started again right away. Other repetitions
were nonadjacent – once the repeating unit ended, other material intervened before the unit returned. Looking at cases where the repetition occurs immediately (without intervening material) can illuminate the relationship between LRU and repetition detection, since the nature of the intervening material is controlled for; there simply is none. Because there were fewer cases where repetition was immediate, we placed LRUs into just two categories: short (LRU < 3.5 s) or long (LRU > 3.5 s). Exposure had a different effect on these two categories of LRU (p = .005), as shown in Figure 5. Specifically, exposure increased the probability of detection for immediate repetitions of long LRUs, but decreased the probability of detection for immediate repetitions of short LRUs. This trend seems to contribute additional evidence for the general theory that exposure helps listeners detect repetitions of long units, but not short units, reflecting a potential shift in attention from low to high hierarchic levels across repeated exposures.

Although the musical contexts of the repetitions in this study were too diverse to be systematically investigated in the logistic regression, there were instances where two repetitions differed by a single primary factor. These isolatable factors were examined with a series of planned pairwise contrasts. For these tests, performance on each of the four exposures of each repetition was examined.

In the Haydn minuet, the first phrase (mm. 1-10) repeats immediately, notated by a repeat sign. But it also returns after the trio, notated by a da capo marking. Despite the strong formal articulation at the return after the trio, the immediate repeat was more noticeable to participants, $\chi^2(1, N = 186) = 40.58$, $p < .001$.

The Haydn provided another interesting opportunity to investigate musical context. The material in m. 8 repeats immediately in m. 9. The material in m. 10 (on its second iteration, at least, within the phrase repeat of mm. 1-10) repeats immediately in m. 11. But phrasal closure intervenes between m. 10 and m. 11, so insofar as the material in the two cases can be considered comparably salient, etc., the two cases isolate the effect of closure on repetition detection. Is it easier to recognize a short, immediate repetition when it occurs within the same phrase or when it starts a new phrase? In the case of the Haydn, the immediate repetition of a measure was more noticeable when it occurred within a phrase rather than between them, $\chi^2(1, N = 216) = 6.60$, $p < .01$.

The opening measures of the Rameau excerpt constitute a kind of parallel period, but articulated over an E pedal. In fact, mm. 5-8 replicate mm. 1-4 exactly, save for a changed cadence figure (the last measure and a half). By the literal definition of this study, this constitutes a 3-measure exact repeat. But this material returns again – mm. 17-24 repeat mm. 1-8 exactly. Given that in the Haydn, the immediate repetition of a phrase was more noticeable than a delayed one, it might be expected that the temporally proximate repetition of the opening material in m. 5 would be more noticeable than the distal repeat in m. 17. But in fact the delayed repeat was more noticeable, $\chi^2(1, N = 214) = 41.65$, $p < .001$. This case differs from the one in the Haydn because the temporally proximate repetition, in m. 5, constitutes a repeat of only part of the initial segment. It could be that the cadence type gets fused so completely with the identity of the material preceding it that people do not conceptualize repeating material as a repetition when the cadence type changes. To recognize something as a repetition, in other words, listeners might need the repeating unit to extend to a point of closure. This is ultimately an empirical issue that could be addressed in a future study with stimuli designed to isolate this factor. Lidov (1979) terms immediate repetitions with altered endings “formative” repetitions – repetitions whose function is to “form” (make evident to listeners) motives and phrases.

The repetition in m. 17-24 repeats the entire first part, including the antecedent and the consequent, and thus embodies within it the same type of three-measure
repetition in the opening (mm. 21-23 repeating mm. 17-20). But this consequent repetition was more noticeable within the m. 17 statement of the theme than in the initial m. 1 statement, $\chi^2(1, N = 191) = 29.05, p < .001$. There are several factors that might account for this. It could be that people are more likely to fuse the antecedent and consequent together at the opening of a piece, to treat it mentally as a single “theme,” and remain blind to its internal structuring, than they are when this theme recurs later in the music. It could be that the repetition some people registered in m. 21 was the four-measure repetition of m. 5, rather than the repetition of m. 17, and the three-measure segment that had occurred in m. 1, 5, and 17. When people described the repetition to which they were responding, they emphasized the salient opening gesture. To really understand what was driving this response, a follow-up study targeting this specific issue would be necessary.

The opening eight bars repeat not only in m. 17, but also in m. 35. The first full repetition of this material, in m. 17, was more noticeable than the second in m. 35, $\chi^2(1, N = 191) = 29.05, p < .001$. There are at least three possible explanations for this difference. First, and least interestingly, respondent fatigue – perhaps participants simply paid less and less attention across the course of each excerpt as boredom set in. Second, it is possible that listeners referred back to the original template (mm. 1-8, in this case) at each repeat, and since more time had passed between that statement and the one in m. 35, it was less salient. Third, musical context may have been playing a more subtle role here. Immediately before the repetition in m. 17, the music does not refer to the opening eighth-note figure; in fact, it uses distinctive quarter-note figures that contrast clearly with the motives from the opening. But the material leading up to m. 35 plays again and again with the opening eighth-note figure, in its original registral position. It is possible that this tautness was over-successful, in the sense that listeners were “fooled” by the repeat in m. 35, not recognizing that the music had stopped playing with that figure and actually employed its original version until it was too late to respond with a successful repetition detection.

Like the Rameau, the Schumann begins with a period structure, but it is stretched out to 16 bars. Also like in the Rameau, this entire opening period returns later in the piece – m. 29 in this case. The results in both cases were similar, the temporally distant repetition of the entire opening phrase was more noticeable than the more immediate repetition (m. 9 in the Schumann) of the opening three measures of the antecedent, $\chi^2(1, N = 209) = 4.05, p < .05$.

Discussion

This preliminary investigation of repetition detection in music demonstrates that not all musical repetitions are created equal. The length of the repeating unit seems particularly important. In particular, in the excerpts investigated in this study, repetition detection was optimal for repeating units of around 6 s.

The contrasts suggest that the repetition detection method could be used in experiments with more rigidly controlled stimuli to reveal more about the effects of musical context. The contrasts in this study imply that immediate, within-phrase repetition is more noticeable than separated, between-phrase repetition, even when the latter is highly formally marked. Moreover, repetition of units that form a complete segment at some level seems to be more noticeable than repetition of units that do not – for example, the kind of partial repetition that characterizes the antecedent and consequent in a parallel period. It’s not that the unit of repetition in this kind of structure is insufficiently long – repetition of measure-level gestures, for example, can be quite noticeable; it’s rather that the repetition is not complete at the level at which the parallelism occurs and, presumably, is conceptualized.

Immediately adjacent repetitions tended to be the easiest to detect, yet a contrast in the Haydn revealed that participants had a hard time identifying an immediate repeat of a measure-long unit when it first functioned as the end of a phrase and then as the beginning. A possible explanation of this result is that the intervening closure made the repetition harder to detect – but in other cases where closure separated the repeating units, such as in the case of the repeat of an entire phrase, participants were good at identifying the repetition. A better explanation, perhaps, is that when it occurred the first time, the measure-long unit was heard as an ending. This ending aspect may have seemed like an essential part of the identity of that unit – it was an ending thing. If this was the case, when the unit recurred – even immediately – as a beginning, it may not have been recognizable as the same object. Follow-up studies that manipulate unit positioning in this way have the potential to reveal interesting information about the perceptual entities, or things (to use Hanninen’s term) that listeners abstract from the ongoing musical stream. In musical conceptualization, the exact notes may be less important than a segment’s function; e.g., whether it serves to begin or end a phrase.

Repeated exposures of the same excerpt modulate the detectability of repeating elements, improving it for
relatively longer units, and worsening it for relatively shorter ones. This pattern of results suggests that attention shifts to progressively higher temporal levels across exposures. The size of the temporal spans to which listeners attend is an important issue in music theory, but has been difficult to address empirically. The repetition detection task employed in this study could potentially be used to investigate the level of temporal spans to which listeners are oriented in different musical contexts. It is particularly promising because it gets at the question of temporal span without asking about it directly; thus circumventing the problem of people’s limited explicit access to this kind of percept while listening.

A serious limitation of this study is the reliance on LRU, a relatively crude measure of musical content. In future studies, stimuli can be selected to isolate more sophisticated musical factors, particularly those that emerged as interesting in the present study’s contrasts. Although this study deliberately used full-textured recordings of real music from several time-periods, future studies might give up a little in ecological validity in order to gain control, by creating synthetic stimuli that isolate particular factors of interest.

This study is further limited by the repertoire used—although the excerpts were selected from divergent corners of common-practice tonal music, it would be interesting to see if repetition functioned similarly in atonal music, or in the highly repetitive contexts of minimalist and process musics. Moreover, this study does not engage the aesthetic consequences of the detectability or lack of detectability of different kinds of repetition; it sometimes seems desirable for a particular repetition to be explicitly recognized as such, but at other times, repetitions seem to purposefully obscure themselves, serving merely to set the stage for some other kind of perception. For example, in the Haydn excerpt used in this study, each line repeats the previous one, but backwards. Although this kind of retroversion was not explicitly asked about in this study, participants did not mention observing it either in the free response portion of the study, nor in the debriefing discussion afterwards. Yet despite its initial lack of immediate detectability, this pattern clearly plays an aesthetic role in the movement. Margulis (in press) investigates the aesthetic consequences of musical repetition.

The question of repetition detectability is but a preliminary step toward a larger understanding of the functions and perceptions of repetition in music. Repetition is uniquely prevalent in music, and uniquely important in theoretical accounts of music. Yet empirical and psychologically oriented studies have yet to devote sufficient attention to this important topic. It is hoped that this study can serve as an impetus to further investigation of the role of repetition, both within and across pieces, in music processing.

**Author Note**

Correspondence concerning this article should be addressed to Elizabeth Hellmuth Margulis, 201 Music Building, University of Arkansas, Fayetteville, AR 72701 E-MAIL: ehm@uark.edu

**References**


