What determines the perception of segmentation in contemporary music?

Michelle Phillips1\*, J. Matthew Wilcoxson2, Emily Howard1, Marcus du Sautoy3, Pip Willcox2, Luke A. Jones4, Andrew J. Stewart4, and David De Roure2 5

1 Royal Northern College of Music, 124 Oxford Road, Manchester, M13 9RD, UK.

2 Oxford e-Research Centre, University of Oxford, Oxford, UK.

3 Mathematical Institute, University of Oxford, Oxford, UK.

4 Division of Neuroscience & Experimental Psychology, University of Manchester, Manchester, UK.

5 The Alan Turing Institute, London, UK.

Correspondence:

Dr. Michelle Phillips

Michelle.Phillips@rncm.ac.uk

Word count (without abstract or references): 5,411 words

ABSTRACT

Background: This study concerns perception of musical segmentation during live listening to contemporary classical music. Little is known about how listeners form judgements of musical segments, although there is some evidence that altering the order of sections of music may not change aesthetic judgements (Eitan & Granot, 2008). Moreover, it is not known how listeners segment music when typical section markers, such as cadences and fermatas, are absent (e.g. Sears, Caplin & McAdams). Contemporary music is often non-tonal and may not prioritize the musical phrase or notion of melody. Aims: The current study aimed to examine which musical features influence listeners’ decisions regarding section endings in a piece of contemporary music, Ligeti’s ‘Fanfares’. Methods: Data were gathered using a smartphone application (‘PRiSM Perception App’), designed for this study by the RNCM Centre for Practice & Research in Science & Music (PRiSM) and the Oxford e-Research Centre. 259 audience participants were asked to ‘tap’ when they felt that a section had ended. Subjective responses were captured, as well as contextual data about the participants. Results: Audience members demonstrated high levels of agreement regarding segmentation. Musical features which influenced decisions included changes in dynamic and changes in register in which motifs were played. Self-report data analysis suggests that listeners were not always aware of how they made decisions regarding segmentation. Conclusions: The factors which may influence judgement of musical segmentation are to some extent similar to those identified by music analysis (Steinitz, 1996) but in other ways different. Trained and untrained listeners exhibit a high degree of agreement on segmentation, even when listening to a piece of music that is unfamiliar to them. Musical training did not significantly impact on judgements of segmentation. However, a listener’s sense of familiarity with contemporary repertoire does seem to influence perception of musical segments – participants who self-reported being familiar with contemporary music marked segmentation more in line with an expert music analysis. This study has implications for contemporary music audiences, as familiarity with contemporary music may have more of an impact on how audiences perceive musical structure than formal musical training.

Keywords: Music perception, large-scale form, segmentation, contemporary music, Ligeti, live music

Introduction

When we listen to a piece of music, we perceive the information as consisting of separate units, or events (Drake & Bertrand, 2001). Such a process may be referred to as grouping (Lerdahl & Jackendoff, 1983), chunking (Miller, 1956), or segmentation (Cambouropoulos, 2006). How listeners infer such chunks of information, or segment the auditory stream, may be influenced by multiple factors. These may include individual differences, such as musical expertise, and/or features in the music itself, for example when the music changes significantly in some way (for example in modality or tempo). Investigating how musical information is parsed in real time, i.e. what determines how one event is deemed to have ended and another begun, may grant valuable insights into the ways in which expertise may shape experience of music and how some aspects in the musical surface or structure may be more perceptually salient than others.

Most popular studies and models of the perception of musical form (e.g. Lerdahl & Jackendoff, 1983; Meyer, 1956), are based on the perception of tonal music. Contemporary (or ‘new’) music is rarely tonal, and hence does not include related features such as diatonic harmony, harmonic tension and release, harmonic closure (e.g. at the cadence), or melody, as conceived in the conventional 19th century sense. Tonal music is often constructed as a series of organised events, existing in a hierarchy. However, this may not be the case for non-tonal music (Dibben, 1994). A common assumption is that “atonal music is generally perceived as opaque and difficult to understand because it rarely makes aural sense to the uninitiated listener” (Spies, 2005). Such a stance raises the question of whether tonal and non-tonal music are perceived differently from one another, and if so, how? Also, to what extent might non-tonal music be more difficult to understand, and does this depend at all on expertise? Research regarding perception of musical form and factors which may influence this (such as a listener’s level of musical training) has only recently begun to interrogate ways in which contemporary classical music is processed and stored (e.g. Schulze, Dowling, & Tillman, 2012), and there are many unanswered questions in this field, despite the obvious relevance to modern day listeners, composers and performers.

Over the last 30 years multiple studies have explored the perception of musical form in tonal music. Most have suggested that local cues (small-scale, short duration) may take priority over global relationships (large-scale, overall formal structure) in the perception of musical information (Deliège, Melen, Stammers, & Cross, 1996, Rolison & Edworthy, 2012; Tillmann & Bigand, 2004). Moreover, studies which have examined the impact of rearranging segments of music on the listener have found that participants may not be able to differentiate between the original and rearranged pieces of music (Tillman & Bigand, 1996), and that listeners may not prefer an original over an arrangement (Eitan & Granot, 2008). Two studies by Granot and Jacobi (2011, 2012) suggested that participants may demonstrate sensitivity to some aspects of musical form (e.g. the ABA structure, the placement of the development section, features at the beginning and end of the work) but not to others (e.g. overall harmonic structure). In summary, this research on perception of large scale form suggests that global relationships in the structure of music may not be available to perception, or if they are, such perception may not be as prominent as the recognition of smaller-scale events.

The question of how such local events are parsed, or segmented, when listening to large-scale form, has only recently begun to be explored. Empirical work has suggested that, in tonal music, cadences (Tillman, Bigand, & Madurell, 1998; Tillman & Bigand, 2004), and long notes and rests (Bruderer, 2008) influence judgement of segmentation during music listening. Studies of non-tonal music are fewer. Clarke and Krumhansl (1990) undertook 5 experiments which investigated various ways in which trained and untrained listeners perceive musical form. Experiments 1-3 used Stockhausen’s *Klavierstueck IX* and Experiments 4-6 used Mozart's *Fantasie in C minor*, K. 475, as their stimulus set. The authors state that, overall, the results of these experiments suggest that attempts to model perception of large-scale form and predict perception of segments (in this case using the model proposed by Lerdahl & Jackendoff, 1983) may be successful, that listeners largely agree on segment boundaries (for both the works by Mozart and Stockhausen), and that listeners are largely accurate in judging location of segments. These results were relevant when listening to both tonal and non-tonal stimuli, which the authors attribute to “the two pieces share[ing] some high-level property on which listeners focus” (p. 249). This study laid important groundwork for questions of perception of musical structure (including segmentation) in tonal and non-tonal music. However, the study also invited a more thorough investigation of the perception of contemporary music (the structure of Stockhausen’s *Klavierstueck IX* may not be representative of non-tonal music more generally, or share attributes with later contemporary music), with a larger participant pool representative of a broader demographic, listening in an ecologically valid environment. The question of how segments are perceived in live performance of contemporary non-tonal music is central to the current study.

Clarke and Krumhansl’s 1990 study used musically trained participants. There is varying opinion regarding whether listeners with musical training perceive musical structure differently from those with no training. Deliege, Melen, Stammers, and Cross (1996) found differences, and Granot and Eitan (2008) found some evidence that musical training may lead to preference of a hybrid version over an original. Ockelford and Sergeant (2012) suggested that musicians may process non-tonal structure differently due to their training; they asked 14 musicians to listen to tone-rows and rate the extent to which notes ‘fit’ with other notes (probe tone paradigm), and found that musically trained listeners may impose tonal frameworks on perception of non-tonal music (however, the study does not include a non-musician group for comparison). In contrast, Tillman and Bigand (2004) found that participants prioritised local over global features in perception, regardless of musical training. Bigand (2002) conducted a series of experiments which explored perception of melody and harmony, and, finding no difference in perception between trained and untrained listeners, concluded that mere exposure to music in everyday life results in everyone being an “experienced listener” (p. 304). A review of studies which examined various ways in which music is processed (perception of tension and relaxation, link between theme and variations, expectation generation, locating local features in global structure, emotional response) by Bigand and Poulin-Charronnat (2006) also found that music perception depends on exposure to music, rather than formal training. Hence studies do not convincingly support claims that musical training may alter perception of musical structure.

The exposure effect (Zajonc, 2001) has been demonstrated to be important in music listening. For example, listeners exposed to a new musical scale for 25-30 minutes may show “extensive learning as characterized by recognition, generalization, and sensitivity to the event frequencies in their given grammar, as well as increased preference for repeated melodies in the new musical system” (Loui, Wessel, & Kam, 2010, p. 377). The proposition in this study, that “knowledge of musical structure is implicitly acquired from passive exposure to acoustical and statistical properties of musical sounds in the environment” (p. 386) is important when considering perception of contemporary music; how this is perceived may depend on the extent to which a listener has been exposed to a similar musical system, or grammar, previously. For example, Western listeners may acquire knowledge of the structure of non-Western music through exposure (Stevens, Tardieu, Dunbar-Hall, Best, & Tillmann, 2013), and both musicians and non-musicians may acquire knowledge of sequences of melodies through exposure to a new musical grammar (Rohrmeier, Rebuschat, & Cross, 2011). On the other hand, studies have suggested that non-tonal music may be more difficult to store in working memory than tonal music, for both musicians and non-musicians (Krumhansl, 1979). Schulze, Dowling, and Tillman (2012) asked musician and non-musician listeners to indicate whether tonal and non-tonal sequences were the same or different to the ones previously heard, and found that both musicians and non-musicians performed better for tonal sequences than non-tonal. These studies suggest that not only may working memory be better for tonal than non-tonal music, but that this may be the case regardless of musical training. It is not clear whether these two theories – that non-tonal music may be processed differently, or that processing depends on exposure – are competing, or two sides of the same coin (i.e. that processing depends on the *extent* of the exposure).

Perception of structure in music relies on the way in which information is grouped and stored in memory, or segmented. Experimental work has recently begun to explore sense of segmentation in contemporary music. Hartmann, Lartillot, and Toiviainen (2017) asked 18 musicians and 18 non-musicians to listen to 6 two-minute stimuli of unfamiliar music, and to note segment boundaries (which they define for participants as “instants of significant musical change”, p. 6) by pressing a computer space bar. In a second experiment the same 18 musicians were invited to listen and mark instants of significant change once again, and then to revisit the places they had marked and move these as they wished after the second hearing. They then rated the strength of each change that they had marked.

The question of what constitutes a “significant change” in music perception is an important one. Models of segmentation have highlighted relevant factors including emotion (Aljanaki, Wiering, & Veltkamp, 2015), rhythm, timbre and harmony (Jensen, 2007), and novelty (Foote, 2000). However, such modelling exercises have largely sought to represent experiences of tonal music (for example, Foote, 2000, discusses a change of tonality as representing a significant change). Most of these musical factors are relevant to contemporary music listening, but few studies have sought to explore how these may lead to perception of change in non-tonal music.

Participants in all studies discussed so far listened to stimuli in a laboratory or office environment. Only recently have studies begun to collect data in a live concert hall setting. Egermann, Pearce, Wiggins, and McAdams (2013) asked 50 concert audience members to provide subjective responses to a live flute performance. In line with theories relating to the exposure effect, they found that listeners’ expectations regarding musical structure were a strong predictor of emotional response. This study provides support for notions such as the ‘experienced listener’ discussed above, i.e. that familiarity with, or exposure to, music may influence perception of segmentation and structure.

Results from relevant literature such as those studies discussed here suggest that there is a need for investigation of the experience of musical form as it unfolds in a live concert hall setting, where much of the real audience listening experience may be preserved. Furthermore, such investigation needs to advance understanding of music created today, rather than tonal music composed in previous centuries. Studies should aim for data gathered from larger audiences, in excess of the 50 participants, which is the maximum in existing studies. Finally, data on environmentally valid listening experiences should be gathered via means which are familiar to listeners, and as unobtrusive as possible on the listening experience. This study seeks to address these gaps in knowledge in this field.

**The current study**

The aims of this study were to examine how audience members perceive segmentation in a live performance of contemporary classical music, and whether perception varies according to musical training, and / or familiarity with contemporary music in general. The study sought to address the gaps in existing empirical work discussed above by employing a smartphone app designed for the study, which allowed live concert hall audience-participants to tap their device when they considered a segment to have ended.

**Research questions**

1. To what extent do participants agree on segment boundaries in a piece of contemporary music?
2. What musical features, if any, occur where there is agreement on segment boundaries?
3. To what extent are there similarities or differences between the structure of the work as perceived, and the structure as identified through music analysis?
4. Does musical training influence decisions of segmentation?
5. Does familiarity with contemporary music influence decisions relating to segmentation?

**Materials and methods**

**Ethical considerations**

The study was granted ethical approval by the Royal Northern College of Music (“RNCM”) Ethics Committee (REC 131, approved 13th September 2017). The first screen of the app contained information relating to ethical consent, and details of how the data would be processed, stored and used.

**Participants**

The 259 participants were audience members who attended an evening concert event in October 2017 as part of the Manchester Science Festival (“MSF”), which is run annually by the Science and Industry Museum in Manchester (“SIM”). The event was advertised in MSF publicity in print and online, and in the event brochure of the RNCM, where the event took place. Those who purchased tickets for the event were subsequently contacted by email and invited to download the PRiSM Perception App (free to download via Apple and Android application online stores) in advance of their attendance at the event. Participation in the study was optional, and audience members could attend the performance without taking part. The 259 audience members who did opt to take part represented 43.2% of the 600 total attendees at the event. The performance was part of the launch event for the PRiSM research centre titled ‘The Music of Proof: What does Maths Sound Like?’ Audience members included musicians from the Royal Northern College of Music, and members (both adults and children) of the general public, with an interest in music, maths, or science communication (as the event was part of a science festival). Participants were not asked for their seat numbers and no data was collected on their physical clustering or dispersal. Peripheral or direct vision of other participants’ activity on the app could have altered behaviour.

Expert musicians were those that reported having had 10 or more years of musical training (52 participants in total). Those familiar with contemporary classical music were considered to be participants who gave a response of 5, 6, or 7 to the question “As a listener, how familiar are you with twentieth-century music?” and those unfamiliar were participants who responded with 1, 2, or 3.

**Apparatus**

Audience responses were captured on participants’ own smartphones or tablets via the PRiSM Perception App, which was designed for this study. The source code used for the mobile app is registered under the DOI “10.5281/zenodo.2542790”. Each device was used by one person and data were captured in real time.

After an ‘about this app’ screen (described above, and including the ethics statements), the app consisted of three pages on which participants were asked to enter data:

1. ‘Your profile’ (date of birth, musical and mathematical training and experience, education, and how often participants listen to music)
2. ‘Performances’ (including the button to tap in response to the live performance)
3. ‘Questions’ (questions relating to experience of the stimulus)

**Stimulus**

Participants heard the solo piano piece ‘Fanfares’ from *Etudes* (Book 1, 1985) by Ligeti. This contemporary study for solo piano was performed in a large concert hall for 600 audience members, of which 259 opted to take part in the current study. The work was selected for the experiment for the following reasons:

1. As a composer, Ligeti is firmly established as representative of contemporary composition
2. Ligeti’s *Etudes* are considered an important part of contemporary music repertoire, and are regularly performed
3. ‘Fanfares’ includes features common to contemporary music (multiple series of notes which repeat, including an ostinato pattern, and a recurring motif commonly referred to as ‘horn fifths’)
4. ‘Fanfares’ lacks many of those commonly found at points considered to mark section boundaries in tonal music (silences, harmonic closure, changes of metre or tempo)

The live performance of this work lasted for 3 minutes and 26 seconds.

**Analysis of ‘Fanfares’ from *Etudes* by Ligeti (from Steinitz, 1996 and personal communication)**

The following is an outline of the structure of ‘Fanfares’ (from Steinitz, 1996 and personal communication):

* **Overall form**:
  + 208 statements of a simple seven-note scale
  + Repetitive ostinati and seemingly spontaneous 'melodies’ (‘horn fifths’)
* **Part 1** (bars 1-45):
  + Bars 1-8: phrase 1 [segment 1]
  + Bar 9: break (silence in one hand) [segment 2]
  + Bar 10-17: phrase 2 [segment 3]
  + Bar 17: break (silence in one hand) [segment 4]
  + Bar 18-26: phrase 3 [segment 5]
  + Bar 27: break (silence in one hand) [segment 6]
  + Bar 28-45: phrase 4 [segment 7]
* **Part 2** (bars 46-115): the melodies increasingly lengthen and merge with each other.  The gaps between the groups of four phrases become shorter and less frequent [segment 8]
* **Part 3** (bars 116-214): Between bar 116 and the end (approx. 95 bars) there is only one gap of 4 quavers, one of 3, and others too short to register as gaps at all [segment 9]

For the purposes of the K-Means cluster analyses detailed below, the above was used to inform the assumption that there were 9 segments in the work.

**Procedure**

Participants could opt to fill in the ‘About You’ page of the smartphone app before they arrived at the performance, or afterwards. Immediately before the performance began, participants were advised to take out their devices and open the ‘Performances’ page on the PRiSM Perception App, which asked them to wait for instruction to begin participation in the study. A countdown was then given to participants by a member of the PRiSM research team, at the end of which they were instructed to tap their screen to synchronise devices to the timings of the performance. The performance then began, and participants tapped the green button displayed on the app whenever they felt that a section had ended. Participants could also tap a button to mark the previous tap as an ‘error’.

Following the experiment, participants were asked to respond to a final series of questions concerning their experience of the piece (on the ‘Questions’ page of the app), including their enjoyment of and familiarity with the piece, familiarity with contemporary music as a genre in general, and how they made decisions regarding section boundaries. A full list of questions can be found in Appendix 1.

**Analysis**

The following steps were used to analyse the data:

1. A Chi squared goodness-of-fit analysis was performed to examine whether total taps were equally spread across the performance, or whether they were not (research question 1)
2. K-Means Cluster analysis was performed to identify the main points in the performance where multiple participants tapped (research question 1)
3. Segmentation per the music analysis of the piece detailed above was compared with absolute number of taps at these moments in the performance (research question 3)
4. The K-Means Cluster analysis above was compared with segmentation as outlined in the music analysis (research question 3)
5. Participants’ self-reports of how they decided where to tap (part of the post-performance ‘Questions’ page of the app) were analysed to determine what the most common factors were in guiding participants’ decisions regarding segmentation (research question 2)
6. Points identified as being the most common for tapping in step 2 above were compared with step 5 above, to ascertain whether people’s self-report of how they made decisions matched where they did indeed tap (research question 2)
7. Step 2 was repeated for musician vs non-musician participants
8. Step 2 was repeated for participants who reported being familiar with contemporary music compared to those who stated that they were not
9. Step 3 was repeated for musician vs non-musician participants
10. Step 3 was repeated for participants who reported being familiar with contemporary music vs those who stated that they were not

**Results**

The 259 participants tapped an average of 8.16 times (standard deviation: 4.43) during the 3.78-minute performance (minimum taps: 1, maximum taps: 21). Taps which participants marked as being ‘errors’ were removed prior to analysis. Figure 1 shows all taps by all participants during the performance of the work, split into one-second bins.

Figure 1. All taps by all participants during the performance of the work, split into one-second bins.

A chi-square test of goodness-of-fit was performed to determine whether the taps were spread evenly over the duration of the performance. The analysis suggested that the taps were not equally distributed across the total duration of the performance, X² (36) = 543.81, p < .0005.

Table 1 shows the time points in the work identified as marking segments according to the music analysis above, and the number of and percentage of total taps for all participants at these points. As can be seen from the table, three of the sections identified by music analysis received a statistically significantly higher number of taps than if the taps were consistently distributed.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Section per music analysis | Bars | Time section ended  (milliseconds) | Number of taps in this one-second interval | |
| Absolute | % of total taps |
| 1 | 1-8 | 13000 | 1 | 0.06% |
| 2 | 9 | 15000 | 4 | 0.25% |
| 3 | 10-16 | 22000 | 4 | 0.25% |
| 4 | 17 | 23000 | 24\* | 1.48% |
| 5 | 18-26 | 31000 | 9 | 0.55% |
| 6 | 27 | 32000 | 33\* | 2.03% |
| 7 | 28-45 | 46000 | 6 | 0.37% |
| 8 | 46-115 | 115000 | 15\* | 0.92% |
| 9 | 116-214 | 212000 | 0 | 0.00% |

Table 1. Time points in the work identified as marking segments according to the music analysis above, and the number of and percentage of total taps for all participants at these points. Asterisked absolute tap numbers indicate where musical segments were noted by both musical analysis and participant identification.

A K-means cluster analysis of all taps, examining the top 9 clusters within the data, gives the clusters shown in Figure 2 (9 clusters were used in the K-means cluster analysis in line with the 9 sections identified by the expert music analyst, as outlined above). The timing of these clusters within the work, and the percentage of total taps which are included in each cluster, is shown in Table 2.

Figure 2. K-means cluster analysis of all taps, examining the top 9 clusters of taps within the data.

|  |  |  |
| --- | --- | --- |
| Cluster | Time (ms) | % of total taps in cluster |
| 1 | 30,192.93 | 21.18% |
| 2 | 81,842.07 | 41.33% |
| 3 | 129,269.04 | 3.10% |
| 4 | 136,387.02 | 6.95% |
| 5 | 155,703.19 | 6.60% |
| 6 | 171,248.07 | 6.26% |
| 7 | 184,582.35 | 1.58% |
| 8 | 192,483.94 | 7.91% |
| 9 | 196,205.99 | 5.02% |

Table 2. The timing (ms) of the 9 clusters within the work, and the percentage of total taps which are included in each cluster.

Cluster 1 coincides with bar 31 in the work. The tapping response seems most likely to have occurred in response to the material in bars 26-27, where there is a break of 13 quavers in one hand of the piano (one hand rests for 13 quavers while the other continues). Cluster 2 coincides with bar 88. This is a moment where multiple aspects of the music change – a break of 10 quavers in one hand is followed by a new rhythmic motif entering in the left hand. The dynamic increases from extremely quiet (volume decreases from ‘pp’) to ‘ff’ suddenly at bar 88.

The cluster analysis of all taps in Figure 2 may also be compared with Table 1, in order to explore whether any of the top clusters of taps matched with key boundaries identified through music analysis. Of those segments identified through music analysis and also marked by participants as being the end of a segment (those marked with \* in Table 1), one of these is also an important cluster: cluster 1 (32,000 milliseconds in Table 1, section 6, and the first cluster point in figure 2 at 30,192 milliseconds). This is the only point where the cluster analysis suggests that all participants marked a segment which coincides with the music analysis. As described above, this is a moment where there is a break of 13 quavers in one hand of the piano.

In the post-performance ‘Questions’ page of the app, participants were asked to state in a blank text box the basis on which they made their decisions of segmentation. Responses were categorised into overarching themes, and the results can be seen in Table 3 below.

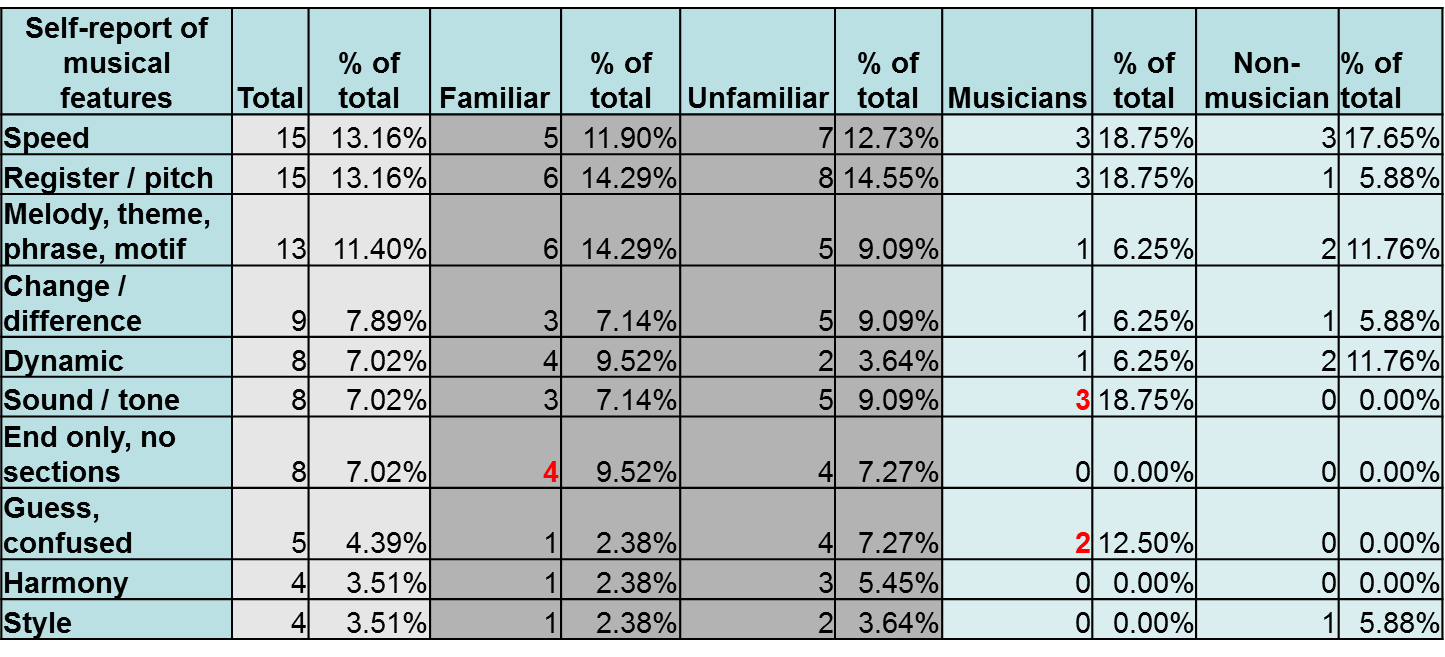


Table 3. Frequency of terms used by audience members regarding how they made decisions as to where a section ended.

As can be seen in Table 3, participants reported that the main factors used to make their decisions about where a segment ended were speed, and pitch (or change of register). This partly aligns with the top clusters in Table 2 – cluster 2 occurs where the dynamic increases from extremely quiet (volume decreases from ‘pp’) to ‘ff’ suddenly at bar 88, and the register of the left hand jumps upwards by more than three octaves. However, these self-reported data do not explain cluster 1 in Table 2, which is marked by a break in the texture of the work – one hand of the piano rests for 13 quavers.

Figure 4 shows a cluster analysis of the top 9 clusters for all participants, musician and non-musician participants. As can be seen in Figure 4, the main clusters (the first and second) were identified by all three groups as being the places at which people tapped most commonly.

Figure 4. K-means cluster analysis of the top 9 clusters for all participants, musician and non-musician participants.

Figure 5 compares the cluster analysis for those familiar and not familiar with contemporary music, with all participants. The timing of these clusters within the work, and the percentage of total taps which are included in each cluster, is shown in Table 4.

Figure 5. K-means cluster analysis of the top 9 clusters for all participants, plus those familiar and not familiar with contemporary music.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Familiar** | | | **Not familiar** | | |
| Cluster | **Time (ms)** | **% of total taps** | **Cluster** | **Time (ms)** | **% of total taps** |
| 1 | 20,299.32 | 18.89% | 1 | 30,117.79 | 17.65% |
| 2 | 59,266.82 | 25.00% | 2 | 64,147.91 | 19.46% |
| 3 | 86,902.67 | 16.67% | 3 | 94,587.00 | 23.08% |
| 4 | 109,634.17 | 6.67% | 4 | 131,100.82 | 9.95% |
| 5 | 133,975.57 | 7.78% | 5 | 145,041.56 | 4.07% |
| 6 | 152,144.60 | 5.56% | 6 | 156,957.14 | 6.33% |
| 7 | 167,105.88 | 4.44% | 7 | 167,506.18 | 4.98% |
| 8 | 182,933.67 | 1.67% | 8 | 181,105.17 | 2.71% |
| 9 | 193,563.65 | 12.78% | 9 | 193,682.72 | 11.31% |

Table 4. Timing of the clusters within the work for those familiar and not familiar with contemporary music, and the percentage of total taps for each group which are included in each cluster.

As can be seen in Table 4, clusters 7-9 were similar for those familiar and not familiar with contemporary music. However, clusters 1-3 demonstrated multiple differences. Firstly, those not familiar with contemporary music tapped at 30,118 milliseconds, which is in line with cluster 1 in Table 2 above (all participants). However, those familiar with contemporary music tapped at 20,299 milliseconds. This point is shortly after bar 17-18 in the piece, where there is a 10 quaver break in one hand of the piano, and the register of both hands moves upwards by the interval of a fourth to a sixth. This aligns with segment 4 in the music analysis in Table 1. Cluster 2 Familiar occurred where there is an 8-quaver break in one hand of the piano, and some change in dynamics and register. Cluster 2 Unfamiliar appears to be a response to the same event, albeit 5 seconds delayed.

**Discussion**

Research questions will be addressed in turn:

1. **To what extent do participants agree on segment boundaries in a piece of contemporary music?**

Results outlined above suggest that taps were not evenly distributed across the piece of music (we would have expected to see a more even distribution, had taps been random). Taken with the clustering of taps, this suggests that there are commonalities in how boundaries are perceived by listeners in this piece.

1. **What musical features, if any, occur where there is agreement on segment boundaries? To what extent are there similarities or differences between the structure of the work as perceived, and the structure as identified through music analysis?**

Clusters 1 and 2 in Table 2 above were the main areas at which all participants tapped. These received 21% and 41% of total taps respectively. The first of these coincides with a segment identified through music analysis, where there is a break in the texture of the piece (a break of 13 quavers in one hand of the piano). However, although three segments outlined in the music analysis did receive a higher number of taps than chance, only one of these segments was one of the main clusters of taps in the whole work. It appears that other factors than those used to inform the music analysis of the piece are more prominent in guiding the judgements of listeners regarding segmentation. Musical characteristics which influence such decisions per the cluster analysis of all participants above include register (pitch), dynamic (volume), and iteration or breakdown of an established pattern (i.e. the greatest number of taps occurred towards the end of the piece, when the continuous ostinato and melody pattern ceases). As detailed above, participants self-reported that speed influenced their decision regarding segmentation. As the work does not change speed (or tempo) throughout, this self-report may be a proxy for a different musical characteristic. Here, perhaps note density has been interpreted as speed (i.e. a point where there are fewer notes in the bar – where one hand of the piano does not play – sounds like a change in tempo).

1. **Does musical training influence decisions of segmentation?**

No. Musical training (participants were considered to be musically trained if they had 10 or more years of musical training) did not appear to influence decisions of segmentation.

1. **Does familiarity with contemporary music influence decisions relating to segmentation?**

Yes, Results suggest that musical experience and expertise have less impact on decisions relating to segment boundaries in contemporary music than familiarity with contemporary music in general. Those familiar with contemporary music tapped more in line with the music analysis (segment 4 in Table 1), and data could be interpreted as this group having reacted sooner to changes in the musical texture. There was therefore more coincidence with the findings of music analysts in the participants who were familiar with contemporary music.

**Conclusion**

Overall, this study suggests that those features identified by a music analyst as marking segment boundaries may not be those which are most salient to music listeners. Although there is some coincidence of the most prominent groups of taps for all participants coinciding with one segment in the music analysis, other clusters occurred at different points. These points seem to have occurred where the music changed significantly in dynamic (volume) or register, or where one hand in the piano part had a rest for 10-13 quavers (NB. Some audience members could see the pianist’s hands and the piano keyboard, but only from a distance). These musical factors which influenced segmentation appear to be evidenced in participants’ self-reports of how they made decisions (providing that change in speed may be interpreted as a change in note density). Findings discussed here therefore support recent discussions of the importance of pitch and dynamics in the perception of musical change. Familiarity with contemporary music seems to have more of an influence on decision of segmentations than musical training. Although there was some overlap in these groups (those with musical training and those familiar with contemporary classical music), the overlap was only partial (46.67% of the participants who responded with 5-7 on a 7-point Likert scale to the question regarding their familiarity with contemporary classical music were also considered to be musicians for the purposes of this study, i.e. they had 10 or more years of musical training). Findings in the current study could be interpreted as supportive of the notion of the ‘experienced listener’.

Future studies should consider conducting audience research in a live concert hall setting, using a smartphone application which can be freely downloaded onto participants’ personal devices. This allows for high resolution and non-obtrusive data capture, which may be conducted in a silent, low-light environment. Data collection by smartphone application also allows for participants to contribute responses both within (real time responses to musical structure) and outside (demographic data and subjective responses to music) the concert hall environment.

Acknowledgements

Many thanks to the Museum of Science and Industry (Manchester) and the Manchester Science Festival team. Thanks to Professor Richard Steinitz for invaluable expertise. This research was partially supported by the programme grant "Fusing Audio and Semantic Technologies for Intelligent Music Production and Consumption" awarded by the Engineering and Physical Sciences Research Council (EP/L019981/1).

**Conflict of Interest Statement**

The submitted work was not carried out in the presence of any personal, professional or financial relationships that could potentially be construed as a conflict of interest.

**References**

Aljanaki, A., Wiering, F., & Veltkamp, R. C. (2015). MediaEval 2015: A Segmentation-based Approach to Continuous Emotion Tracking. Paper presented at the MediaEval.

Bigand, E. (2002, Oct 25-27). More about the musical expertise of musically untrained listeners. Paper presented at the Conference on Neurosciences and Music: Mutual Interactions and Implications on Developmental Functions, Venice, Italy.

Bigand, E., & Poulin-Charronnat, B. (2006). Are we "experienced listeners"? A review of the musical capacities that do not depend on formal musical training. *Cognition*, 100(1), 100-130.

Bruderer, M. J., & McKinney, M. F. (2008). Perceptual evaluation of models for music segmentation. Paper presented at the Proceedings of the 4th Conference on Interdisciplinary Musicology.

Cambouropoulos, E. (2006). Musical parallelism and melodic segmentation: A computational approach. *Music Perception*, 23(3), 249-268.

Clarke, E. F., & Krumhansl, C. L. (1990). Perceiving Musical Time. *Music Perception*, 7(3), 213-252.

Deliège, I., Melen, M., Stammers, D., & Cross, I. (1996). Musical schemata in real-time listening to a piece of music. *Music Perception*, 14(2), 117-159.

Dibben, N. (1994). The cognitive reality of hierarchic structure in tonal and atonal music. *Music Perception*, 12(1), 1-25.

Drake, C., & Bertrand, D. (2001). The quest for universals in temporal processing in music. *PsychoL Sci*, 13, 71-74.

Egermann, H., Pearce, M. T., Wiggins, G. A., & McAdams, S. (2013). Probabilistic models of expectation violation predict psychophysiological emotional responses to live concert music. *Cognitive, Affective, & Behavioral Neuroscience*, 13(3), 533-553.

Eitan, Z., & Granot, R. Y. (2008). Growing Oranges on Mozart's Apple Tree: "Inner Form" and Aesthetic Judgment. *Music Perception*, 25(5), 397-418.

Foote, J. (2000). Automatic audio segmentation using a measure of audio novelty. Paper presented at the 2000 IEEE International Conference on Multimedia and Expo. ICME2000. Proceedings. Latest Advances in the Fast Changing World of Multimedia (Cat. No. 00TH8532).

Granot, R. Y., & Jacoby, N. (2011). Musically Puzzling I: Sensitivity to overall structure in the sonata form? *Musicae Scientiae*, 15(3), 365-386.

Granot, R. Y., & Jacoby, N. (2012). Musically puzzling II: Sensitivity to overall structure in a Haydn E-minor sonata. *Musicae Scientiae*, 16(1), 67-80.

Hartmann, M., Lartillot, O., & Toiviainen, P. (2017). Interaction features for prediction of perceptual segmentation: Effects of musicianship and experimental task. Journal of *New Music Research*, 46(2), 156-174.

Jensen, K. (2007). Multiple scale music segmentation using rhythm, timbre, and harmony. *EURASIP Journal on Applied Signal Processing*, 2007(1), 159-159.

Krumhansl, C. L. (1979). The psychological representation of musical pitch in a tonal context. *Cognitive Psychology*, 11(3), 346-374.

Lerdahl, F., & Jackendoff, R. (1983). A Generative Theory of Tonal Music. Cambridge, Mass; London: MIT Press.

Loui, P., Wessel, D. L., & Kam, C. L. H. (2010). Humans rapidly learn grammatical structure in a new musical scale. *Music Perception*, 27(5), 377-388.

Meyer, L. B. (1956). Emotion and Meaning in Music. Chicago: Chicago University Press

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*, 63(2), 81-97.

Ockelford, A., & Sergeant, D. (2012). Musical expectancy in atonal contexts: Musicans' perception of "antistructure". *Psychology of Music*, 41(2), 139-174.

Rohrmeier, M., Rebuschat, P., & Cross, I. (2011). Incidental and online learning of melodic structure. *Consciousness and Cognition*, 20(2), 214-222.

Rolison, J. J., & Edworthy, J. (2012). The role of formal structure in liking for popular music. Music Perception: An Interdisciplinary Journal, 29(3), 269-284.

Schulze, K., Dowling, W. J., & Tillmann, B. (2012). Working memory for tonal and atonal sequences during a forward and a backward recognition task. *Music Perception*, 29(3), 255-267.

Sears, D., Caplin, W. E., & McAdams, S. (2014). Perceiving the classical cadence. *Music Perception*, 31(5), 397-417.

Spies, B. (2005). Facilitating access to atonal music: Ligeti's second string quartet. *Journal of the Musical Arts in Africa*, 2, 55-69.

Steinitz, R. (1996). The Dynamics of Disorder. *The Musical Times*, 137(1839), 7-14.

Stevens, C. J., Tardieu, J., Dunbar-Hall, P., Best, C. T., & Tillmann, B. (2013). Expectations in culturally unfamiliar music: influences of proximal and distal cues and timbral characteristics. *Frontiers in Psychology*, 4, 789.

Tillmann, B., & Bigand, E. (1996). Does Formal Musical Structure Affect Perception of Musical Expressiveness? *Psychology of Music*, 24, 3-17.

Tillmann, B., & Bigand, E. (2004). The Relative Importance of Local and Global Structures in Music Perception. *The Journal of Aesthetics and Art Criticism*, 62(2), 211-222.

Tillmann, B., Bigand, E., & Madurell, F. (1998). Local versus global processing of harmonic cadences in the solution of musical puzzles. *Psychological Research*, 61(3), 15-174.

Zajonc, R. B. (2001). Mere exposure: A gateway to the subliminal. *Current Directions in Psychological Science*, 10(6), 224-228.

**Appendix 1 – Survey questions**

|  |
| --- |
| Q1. What is your date of birth (YYYY-MM-DD)? |
|  |
| Q2. How many years of formal music training have you had (including A-level and any instrumental, vocal or composition lessons)? [options: 0 1 2 3 4 5 6 7 8 9 10+] |
|  |
| Q3. Do you currently play a musical instrument, sing or compose, and if so for how long? (Please select zero if you do not have any) [options: 0 1 2 3 4 5 6 7 8 9 10+] |
|  |
| Q4. How many years of formal mathematics training have you had (including A-level and any further study of mathematics)? [options: 0 1 2 3 4 5 6 7 8 9 10+] |
|  |
| Q5. Do you currently work in a field which requires mathematical skills, and if so for how long have you worked in this area? (Please select zero if you do not work with mathematics) [options: 0 1 2 3 4 5 6 7 8 9 10+] |
|  |
| Q6. What is your highest level of formal qualification? [options: GCSEs, A-levels, Bachelor's Degree, Masters Degree, PhD, other (please state in text box)] |
|  |
| Q7. How often do you listen to music (of any style)? [options: never, occasionally, sometimes, most days, every day] |
|  |
| Q8. How long did you think that the piece of music lasted? [minutes and a seconds box, with the options that either can be left blank but not both] |
|  |
| Q9. Please describe the piece in three words [followed by text box] |
|  |
| Q10. How did you decide when a section had ended [followed by text box]? |
|  |
| Q11. How much did you enjoy the piece of music? (from 1 = I didn't enjoy it at all to 7 = I enjoyed it a lot) [options 1-7] |
|  |
| Q12. As a listener, how familiar are you with twentieth-century classical music? (from 1 = I am not familiar with it at all to 7 = I am very familiar with it) [options 1-7] |
|  |
| Q13. How often do you listen to twentieth-century classical music? (from 1 = never listen to 7 = listen every day) [options 1-7] |
|  |
| Q14. How familiar are you with the piece performed? (from 1 = I’ve never heard of it to 7 = I’ve heard it many times) [options 1-7] |
|  |
| Q15. Does participation in a scientific experiment such as this increase or decrease your enjoyment of a performance? (from 1 = it significantly decreases my enjoyment to 7 = it significantly increases my enjoyment) [options 1-7] |
|  |
| Q14. What motivated you to come to tonight's event (select all that apply)? [options: I wanted to learn more about music and maths working together, I am a regular attendee of RNCM events, I wanted to take part in a scientific study, a friend / family member asked me to come along] |
|  |
| Q16. Please use this box for any other comments you wish to make [followed by text box] |