

## Deployments of change and novelty in a corpus of popular music

Christopher Wm. White<sup>1†</sup>

<sup>1</sup> University of Massachusetts Amherst, Amherst, MA, USA

† Corresponding author: [cwwhite@umass.edu](mailto:cwwhite@umass.edu)

Published 16 December 2021; <https://doi.org/10.18061/FDMC.2021.0014>

Author video presentation and/or other conference material: <https://doi.org/10.17605/OSF.IO/MPBT5>

### Abstract

This paper considers aspects of form and texture in 20th-century American popular music using a corpus of textural/lyrical/vocal annotations added to the McGill-Billboard corpus (Burgoyne, 2012). The analysis tracks two categories of events: 1) *changes* differ from immediately prior events; while 2) *novelties* have not happened previously within a song. I additionally divide these events into their specific species of change (i.e., whether the change involves the introduction of new instrumentation, new harmonizations, etc.), and correlate these events with their position in each track in terms of both clock time and formal zone. These interconnections allow me to demonstrate the types of events that appear at particular places within a song, and suggest ways that listeners may use musical cues to form expectations about a song's form and length.

KEYWORDS: *corpus analysis, popular music, lyrics, timbre, texture, expectation*

### Introduction

One utility of large-scale musical forms is their ability to orient a listener within the timeline of a musical event (Hepokoski and Darcy, 2006). For instance, a listener familiar with the norms of 18<sup>th</sup> century Viennese sonata forms will know that, upon hearing the beginning of the piece's development section, that the movement is roughly at its halfway point.

20<sup>th</sup>-century American popular music, especially that based on the rotation of the verse/chorus pair (Margulis, 2014; Temperley 2018), poses a problem for such formal wayfinding: forms that repeat the same musical materials in rotation pose a potential difficulty for a listener to orient themselves, given that the events of the beginning, middle, and end share much of the same musical material (Brody, 2020; Puri, 2020).

This project uses corpus analyses to demonstrate ways that certain musical invariants – specifically, textural, timbral, lyrical, rhythmic, and harmonic changes – are consistently associated with certain points in the timelines of American popular songs. This work combines the notion that studies in form should be grounded in the norms of some corpus (Brody, 2017; Greenberg, 2017; Hepokoski & Darcy, 2006) with the

idea that texture, timbre, and repetition play fundamental roles in popular music's construction (Barna, 2019; deClercq & Temperley 2011; Lavengood, 2017; Temperley, 2018). To this end, this study analyzes the McGill-Billboard corpus (a dataset of chord and formal annotations of American popular music, 1958-1991; Burgoyne, 2012) supplemented with the textural annotations of White et al. (2021) to show that certain events can be associated with particular points within a popular song's form. In the end, I suggest that these formal associations may provide a way for listeners to orient themselves within the rotations and repetition that characterize American popular music.

### General Method

#### *The Corpus*

For this analysis, I use the McGill Billboard Corpus (Burgoyne, 2012), a corpus that spans 1958 to 1991, and uses songs selected from the Billboard Top 100 for each year. The corpus itself consists of text files indicating the song's title, key, likely time signature, and finally lines of text corresponding to each musical phrase. These lines indicate the phrases' chords in an adapted lead-sheet-style format, along with measure delineations, designations for the beginning of formal zones, and micro-timing stamps for the phrase beginnings. Further annotations have been added to 665 of these files (White, et al. 2021). These annotations correspond to moments of change within the texture, timbre, text, or other sonic parameters of the song. The annotations each have three values, with the first value representing the category of change (indicating the musical domain or instrument group corresponding to the change, like “percussion” or “voice”), the second adding specificity to the change (e.g., specifying “backup singer” within the broader “voice” category), and the third indicating the type of event (either “change,” “entry,” or “exit”). These annotations use 50 unique three-character annotations, with 14,376 total annotations.

#### *Methodologies*

Given that this project's motivation is to identify general trends within a dataset, an exploratory (rather than



experimental) stance was adopted. Three main properties of this dataset were examined. First, information concerning the clock-time length of each track was compiled, along with the annotations’ *relative placement* within this clock-time. The latter of these values is calculated as the proportion between an event’s clock time and the overall clock-time duration of the track (i.e., the relative placement of an annotation that appears one minute into a four-minute song would be .25). Second, changes indicated by the corpus’s annotations were also associated with formal zones as indicated by the original corpus. Third, “novel” events – event types that were yet to occur within the track – were tracked both in connection to their position in proportional timelines and in regard to formal zones.

### Observations

#### Clock Time

Figure 1 shows the distribution of the clock-time duration of songs within this corpus. The distribution shows the influence of the “three minute rule” in popular music, a norm that arose from the limitations of early recording technologies in which each side of a record could encode between roughly 3 minutes (for a 10” 78 RPM recording) and 5 minutes (for a 12” 78 RPM recording; Mudge & Hoek, 2001) of material. Indeed, 90% of songs within this dataset clock between 2.2 and 4.8 minutes. Such a constraint upon the length of a song is well documented within popular music literature (McKinney, 2015; Temperley 2018). However, it is an important baseline consideration when building a notion of form within this repertoire.

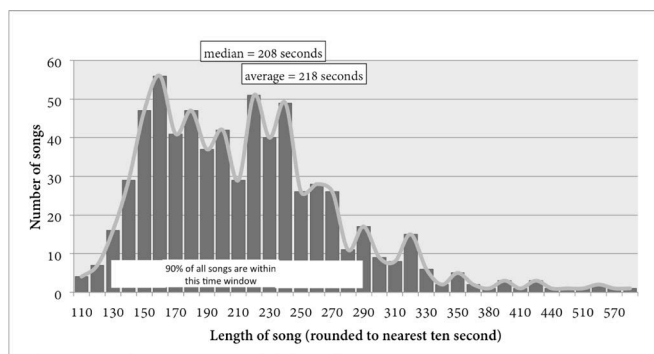


Figure 1: Clock-time duration of songs within the corpus

#### Changes within a Timeline

Table 1 shows twelve of the corpus’s most salient annotations and associates them with the average position at which they tend to appear within a track.

Positioning is shown in two ways. The *mean* value simply averages the relative placements of each annotation; the standard deviation (SD) and number of observations (*n*) associated with this mean are provided as well. Given that these distributions often do not all approximate a normal – let alone a unimodal – distribution, “spikes” within the distributions are also identified. Events’ relative placements were rounded to the closest single decimal, and distributions of those rounded values are created for each event type. “Spikes” were identified as points in these distributions that peaked 10% or more from surrounding values; if two consecutive values remained high, they were considered part of the same peak. Table 1’s leftmost column shows the three approaches used to identify salient annotations: 1) identifying the three most-frequent annotations, 2) identifying the three most-consistent annotations (i.e., those with the lowest SD values), and finally identifying those six further annotations that maximize frequency and consistency (approximated by the lowest standard error values, excluding event types that appear in the prior two categories).

Table 1: Associations between event-types and position within a track

|  | Type                              | Spike-points (% of track) | Mean point in timeline: | SD: | <i>n</i> |
|--|-----------------------------------|---------------------------|-------------------------|-----|----------|
| Most Frequent                            | New Lyrics                        | 0-10%                     | .34                     | .27 | 4493     |
|  | New Singer Enters                 | 0-20%, 60-80%             | .33                     | .28 | 1296     |
|  | Accompaniment instrument enters   | 0-10%                     | .28                     | .27 | 990      |
| Most Consistent (lowest SD)              | Fadeout                           | 80-100%                   | .91                     | .06 | 283      |
|  | Repeated lyrics set in new way    | 90-100%                   | .8                      | .14 | 70       |
|  | Key change                        | 50-80%                    | .56                     | .19 | 119      |
| Most Frequent and Consistent (lowest SE) | Entry of backup singers           | 20-40%                    | .42                     | .25 | 958      |
|  | Exit of lead singer               | 40-60%, 90-100%           | .54                     | .25 | 778      |
|  | Exit of backup singers            | 30-60%                    | .47                     | .22 | 582      |
|  | Enter solo instrument             | 0-10%, 50-70%             | .34                     | .26 | 521      |
|  | Full percussion section enters    | 0-10%                     | .2                      | .27 | 458      |
|  | Full instrumental ensemble enters | 0-10%                     | .2                      | .29 | 488      |

Two main observations arise from this table. First, there are a handful of events that most often appear at either the very beginning or end of tracks, with the entry of new lyrics, singers, percussion, and accompaniment textures instantiating the former, and fadeouts and exits of singers instantiating the latter. On the one hand, this observation may seem trivial: *of course* these types of

events begin these tracks. After all, in order for a song to begin, singers and instrumentalists need to begin playing and singing. However, within this repertoire, such observations are not trivial. With the exception of live concerts, these songs are presented within seamless sonic streams often with little or no silence in between (consider a progression of songs played on the radio, the sequence of a playlist, the background music of an exercise class, or a wedding DJ’s setlist). These cues, therefore, are important to unambiguously indicate that a new form has begun, and that the prior song is complete – they define what it means to *begin* a popular song’s form. Second, several of these events occur at particular intermediate points within a song’s timeline. For instance, the entry of a new singer or solo instrument most often occurs 60-80% or 50-70% through a track, respectively. Such events potentially serve as markers within these songs’ overall forms.

Table 2 shows the annotations most often associated with each formal zone used in the McGill-Billboard corpus. To compensate for the different frequencies of each formal zone, I recorded annotations whose percentage of entries met or exceeded 1.5 times the percentage with which a formal zone occurred. (For instance: 22.8% of all formal zones are choruses: for an annotation to be associated with this zone, more than 34.2% of the annotation’s occurrences had to occur within that zone.) Only events that occurred more than 50 times in the dataset were considered. Events that are shared by zones are italicized and color-coded.

Table 2: Associations between event-types and formal zones

| Verse   | Interlude   | Outro   | Intro   |
|---|---|---|---|
| - Entry of lead singer<br>- Call/response or echo format begins<br>- New Lyrics   | - Rhythmic density decreases<br>- Percussion solo/break begins/ends | - Repeated lyrics set in new way<br>- Fadeout   | - Enter full ensemble<br>- Enter full percussion<br>- Enter accompaniment instrument<br>- Enter solo instrument   |
| Chorus  |   |   |   |
| - Entry of backup singers<br>- Choir entry/exit<br>- Fadeout<br>- Full percussion enters<br>- Key change<br>- Enter/exit full ensemble<br>- Overall range expands |   | - Rhythmic density increases<br>- Call/response format begins and ends<br>- Begin new vocal register<br>- New vocal or harmonization technique used<br>- Repeated lyrics set in new way |   |
| Bridge  | Solo  | Pre-chorus  | Transition  |
| - Key change<br>- Accompaniment ensemble changes<br>- Spoken lyrics begin   | - Enter/exit solo instrument  | - Accompaniment ensemble changes<br>- Entry/exit of backup singers<br>- Add digital sounds to acoustic texture<br>- Begin new vocal register  | - Overall range contracting<br>- Rhythmic density decreases<br>- Key change<br>- New Lyrics<br>- Exit lead singer |

The table illustrates several notable features of this corpus. First, only a handful of event types are *not* shared between formal zones (i.e., the black, non-italicized events). The presence of these events would seem to uniquely point to a particular formal zone. For instance, a listener hearing the overall range of the song’s texture expand could reasonably conclude that they are hearing a chorus. Second, each of these zones has a unique slate of associated events. Such groups of events may contribute to these zone’s definitions and identity. For instance, part of the definition of “chorus” may be the inclusion of the event types included in Table 2. Similarly, these zones can also be defined by what they do *not* feature: part of the identity of a chorus may be that it does *not* feature the events associated with other zones.

However, as indicated by the colored italics of Table 2, there exists a reasonable amount of overlap between zones. Table 3 undertakes a *k*-means cluster of zones using normalized frequencies of all event types that occur greater than 50 times in the corpus. (Using an analysis of silhouette widths for  $k = [1...8]$ ,  $k=4$  was shown as the optimal clustering; White and Quinn, 2018.) The table shows the families of zones that are associated with similar event types with similar frequencies, along with a brief description of the sorts of events types uses in that cluster (i.e., the “character”). Finally, the “secondary cluster” shows the cluster to which the zone is next closest, affinities which illustrate secondary characteristics of each zone.

Table 3: *k*-means clustering of each zone’s event-type associations ( $k=4$ )

|              | cluster | “Character”   | secondary cluster |
|--------------|---------|---|-------------------|
| bridge       | 1       | Many changes, decreases in energy, exits                                    | 3                 |
| trans        | 1       |   | 3                 |
| verse        | 1       |   | 2                 |
| prechorus    | 2       | Increased thickness of textures range increases, new entries                | 1                 |
| chorus       | 2       |   | 3                 |
| instrumental | 3       | Decreases in rhythm and texture, entries and exits of non-vocal instruments | 1                 |
| solo         | 3       |   | 1                 |
| interlude    | 3       |   | 1                 |
| pre-verse    | 3       |   | 4                 |
| outro        | 3       |   | 1                 |
| intro        | 4       | Entries   | 1                 |

*Disposition of Novelty*

Figure 2 shows the percentage of annotated events that have not occurred within the song before (i.e., “novelties”) versus events that have occurred before, again using relative placement values rounded to the nearest 10%. Naturally, the initial events of a song are mostly novel – there has not yet been an opportunity to repeat these events. As the song continues, fewer and fewer of the annotations are novel: unsurprisingly, as verses and choruses repeat, they also repeat the same types of events. However, the last moments of these songs feature an uptick in novelty – the last several moments of a song tend to issue musical changes unlike those that have occurred previously. A listener who hears an increase in novel events after a period of repetition might be cued to a song’s impending ending (at least if they are attuned to this particular norm of this repertoire).

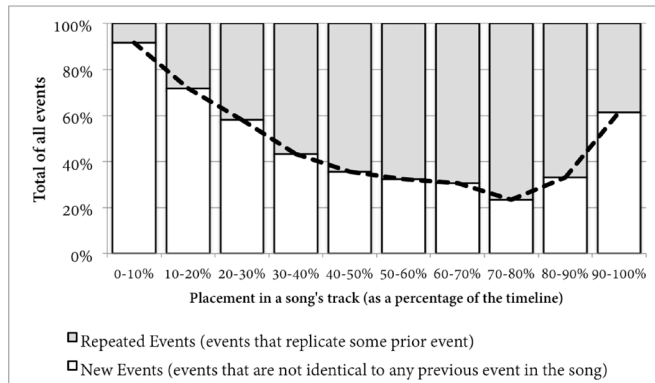


Figure 2: Use of “Novelty” within songs’ timelines

*Disposition of Formal Zones and the Interaction of Parameters*

Figure 3 shows the overall distribution of formal zones within this corpus, with choruses and verses dominating the graph with nearly 60% of the mass. As has been observed by many observers (and argued in Temperley, 2018), these two zones act as a formal backbone in this repertoire. Indeed, every song in this corpus has at least one of these zones. Due to their frequency and ubiquity, we can imagine these zones acting as *core modules* within this repertoire: each song features one or both of the chorus/verse zones, and the repetition and rotation of these modules is a central feature of the song’s form.

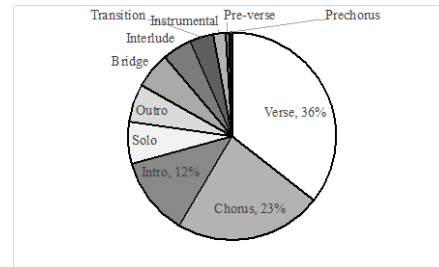


Figure 3: Distribution of formal zones

Figure 4 shows how these core modules are deployed, how they interact with other formal zones, and how other musical parameters are used alongside these zones. The top of the figure presents a flowchart of probabilities derived from the corpus, showing a series of decisions concerning the construction of core modules, how often to rotate the core, and what other formal zones might be included within the song. Clock-time duration is shown above the flowchart, and a proportional timeline is shown below. Finally, a selection of musical changes are shown below the proportional timeline, with their position indicating approximate positions within the track, and percentages again indicating how often such events appear. Overall, this figure provides a birds-eye view of how formal zones are ordered and rotated within this repertoire, and how these orderings interact with certain musical and lyrical changes.

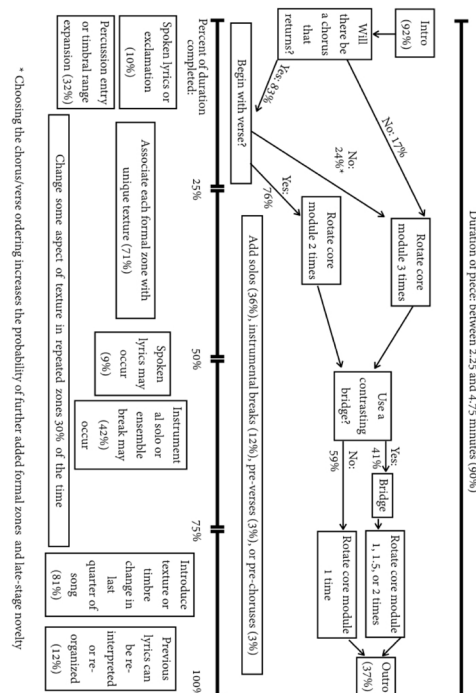


Figure 4: Disposition of formal zones and selected changes within an overall timeline

## Discussion and Conclusion

Overall, this study suggests four strategies for articulating the progression of a form within a repertoire of American popular songs. First, clock time presents an important overall constraint, with most songs lasting somewhere between 2.2 and 4.8 minutes. Second, certain types of events are often associated either with particular locations within a track, or are associated with particular formal zones. Third, novel events (types of events that have not yet been heard within a track) are clustered at the beginning of a track, and again at the end, with intervening points repeating prior event types. Finally, there exist certain norms concerning the deployment and rotation of formal zones within a song.

This paper began by suggesting that the repetition inherent in American popular music poses a potential problem for listeners: it is difficult to know one's position in a timeline if the song consists primarily of repetitions of the same musical material. However, this study has shown several musical events that have a normative placement within songs' overall durations, and that these events can even be uniquely associated with particular formal zones within this repertoire. These norms may be available to listeners when they are creating expectations about a song's form, and if they are sufficiently internalized can help orient listeners within a song's overall timeline. Furthermore, such norms are based not solely on melodic and harmonic parameters, but on textural, timbral, lyrical, and rhythmic variations. Indeed, this suggests that these domains – and the variation contained therein – may play crucial roles in listener's understanding of form in 20<sup>th</sup>-century American popular music.

## References

- Barna, A. 2019. *Examining contrast in rock and popular music*. Ph.D. Dissertation, Eastman University.
- Brody, C. (2016). Parametric interaction in tonal repertoires. *Journal of Music Theory*, 60 (2), 97–148. <https://doi.org/10.1215/00222909-3651814>
- Brody, C. (2020). Two Langerian sonata-Form problems, with solutions by Beach and Medtner. *Paper delivered at Society for Music Theory Annual Meeting*.
- Burgoyne, J.A. (2012). *Stochastic processes and dataset-driven musicology*. Ph.D. dissertation, McGill University.
- Byros, V. (2009). *Foundations of Tonality as Situated Cognition, 1730–1830*. Ph.D. Dissertation, Yale University.
- deClercq, T, & Temperley, D. (2011). A corpus analysis of rock harmony. *Popular Music*, 30(1), 47–70. <https://doi.org/10.1017/S026114301000067X>
- Greenberg, Y. (2017). Of beginnings and ends: A corpus-based inquiry into the rise of the recapitulation. *Journal of Music Theory*, 61 (2), 171-200. <https://doi.org/10.1215/00222909-4149546>
- Hepokoski, J & Darcy, W. (2006). *Elements of Sonata Theory: Norms, Types, and Deformations in the Late-Eighteenth-Century Sonata*. New York: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195146400.001.0001>
- Lavengood, M. (2017). A new approach to the analysis of musical timbre. Ph.D. diss, CUNY Graduate Center.
- Margulis, L. (2014). *On repeat: how music plays the mind*. New York: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199990825.001.0001>
- McKinney, K. (2015). A hit song is usually 3 to 5 minutes. here's why. *Vox*. [vox.com/2014/8/18/6003271/why-are-songs-3-minutes-long](http://vox.com/2014/8/18/6003271/why-are-songs-3-minutes-long)
- Mudge, S., D.J. Hoek. (2001). Describing jazz, blues, and popular 78 RPM sound recordings: suggestions and guidelines. *Cataloging & Classification Quarterly*, 29(3). 21-48. [https://doi.org/10.1300/J104v29n03\\_03](https://doi.org/10.1300/J104v29n03_03)
- Puri, M. (2020). Rituals of circularity: on the conceptual underpinnings of rotational form.” *Paper delivered at the Society for Music Theory Annual Meeting*.
- Riley, M. (2010). Hermeneutics and the new Formenlehre: an interpretation of Haydn's 'Oxford' symphony, first movement. *Eighteenth-Century Music*, 7(2): 199 – 219. <https://doi.org/10.1017/S1478570610000047>
- Temperley, D. (2018). *The musical language of rock*. New York: Oxford University Press. <https://doi.org/10.1093/oso/9780190653774.001.0001>
- White, C. and Quinn, I. (2018). Chord content and harmonic function in tonal music. *Music theory spectrum*, 40(2), 314-350. <https://doi.org/10.1093/mts/mtv021>
- White, CW., Fulmer, J., Cordova, B., Black, A., Danitz, C., Evans, W., Fischer, A., Greene, R., He, J., Kenyon, E., Miller, J., Moylan, M., Ring, A., Schwitzgebel, E., & Wang, Y. (2021). A New Corpus of Texture, Timbre, and Change in 20th-Century American Popular Music. in *Proceedings of the Future Directions of Music Cognition International Conference, 6–7 March 2021*. <https://doi.org/10.18061/FDMC.2021.0015>