The Decline of Harmonic Schemata in Popular Music Chord Loops

Aditya Chander and Ian Quinn

May 30, 2023

The decline of harmonic schemata in popular music chord loops

Aditya Chander (a), Ian Quinn (b)

(a) Yale University, USA, aditya.chander@yale.edu, (b) Yale University, USA, ian.quinn@yale.edu
*corresponding author

Several accounts of popular harmony focus on schemata that repeat across songs. In contrast, de Clercq & Margulis (2018) note that there is considerable stylistic variety in popular music, suggesting that listeners rely more on dynamic expectations for harmony (guided by extensive repetition within songs) than schematic expectations (from patterns occurring across songs). However, they do not situate their claims in a specific era of popular music. Thus, we perform an analysis using tools from information theory and natural language processing to predict how frequently short-term expectations may be engaged in two popular music corpora from different time periods: the McGill Billboard corpus (1958–1991) and the Yale-Geerdes Billboard Corpus (2002–2021). Three key results emerge from this analysis. Firstly, the Yale-Geerdes corpus is significantly more repetitive than the McGill corpus. Secondly, a large majority of chord loops in each corpus appear in only one song, but the most schematic loops in the McGill corpus appeared in more songs than the most schematic loops in the Yale-Geerdes corpus. Thirdly, the average repetitiveness of loops appearing in one song was also significantly higher in the Yale-Geerdes than the McGill corpus. Much twenty-first-century popular music thus affords the engagement of dynamically adapting expectations for distinctive harmonies.

Keywords: popular music, harmony, corpus studies, loops, expectation

1. Introduction

1.1. Background

Several accounts of popular music harmony recognise *schemata* – progressions of 2-8 chords that recur across several songs – as the building blocks of musical form (Acevedo, 2020; Doll, 2017; Nobile, 2020; Tagg, 2014). While extensive catalogues of schemata have been developed to account for goal-directed tonal motion (Biamonte, 2010; Doll, 2017; Nobile, 2020), schematic patterns are also used frequently as the basis for highly repetitive looping passages (Tagg, 2014), where harmonic goal-directedness is suspended in service of forming a backdrop for other musical parameters such as melody. Importantly, musical schemata have been framed in psychological terms as mental templates that guide a listener's perception of musical structure and underpin their expectations (Acevedo, 2020; Bharucha, 1987; Gjerdingen, 2007).

While some studies use close listening and manual analysis to catalogue schemata from hundreds or even thousands of songs (Doll, 2017; Duinker, 2019; Richards, 2017; Tagg, 2014), the increasing availability of symbolically encoded popular music corpora has resulted in a number of computational studies of schema identification (Acevedo, 2020; de Clercq & Temperley, 2011; Sears & Forrest, 2021). These computational studies directly model the psychological mechanism of statistical learning (Aslin, 2017), where chord frequencies and transition probabilities are acquired automatically from long-term exposure to a musical style. Statistical learning has been shown to underpin the learning of novel musical grammars (Jonaitis & Saffran, 2009; Loui et al., 2010) and explains listeners' note-by-note pitch expectations for Bach chorale melodies (Pearce, 2018). Under the statistical learning account, schemata are chord progressions within the style that have high transition probabilities from one constituent chord to the next.

However, long-term schematic expectations are not the only types of expectation that are available to a listener. Huron (2006) notes that listeners may also use *dynamic expectations* to adapt to the incoming musical information. In the realm of popular music studies, de Clercq and Margulis (2018) suggest that dynamic expectations are particularly powerful as popular music features a high degree of internal repetition and looping. Indeed, they claim that schematic expectations might not be robust in popular styles as there is considerable variety in the harmonic logic of popular music. However, they do not provide examples of non-schematic loops nor a method for finding them. Additionally, they do not situate their claim within a particular style or period of popular music.

1.2. Aims

In this study, we sought to add nuance to de Clercq's and Margulis's claims by exploring whether they apply more to harmony in 20th- or 21st-century popular music. We report an analysis of two popular music corpora in which we measure two related properties. We estimate the *distinctiveness* of chord loops both at the level of a song and at the level of the corpus. These different levels of inquiry change the definition of distinctiveness: a loop that is distinctive in the context of a song is *repetitive* within that song; by contrast, a loop that is distinctive in the context of a corpus appears *rarely*. These analyses predict the extent to which dynamic expectations may be engaged in popular music from this century and the last.

1.3. Term frequency (TF) and inverse document frequency (IDF)

To quantify distinctiveness, we used two metrics developed for natural language processing applications: the *term frequency* (hereafter *TF*; Luhn, 1957) and the *inverse document frequency* (hereafter

IDF; Spärck Jones, 1972). The TF is the proportion of a document made up of a term (usually a word); in other words, it is a measure of repetitiveness. Meanwhile, the IDF is the log inverse proportion of documents in a corpus containing a term; in other words, it is a measure of rarity. We define these mathematically in Section 2.4.

The TF and IDF are often multiplied to create a single metric, the TFIDF, which represents the importance of a term to a document. In the music information retrieval literature, TFIDF scores have been used to weight different musical features in service of some other task, such as classifying the mode of plainchant (Cornelissen et al., 2020) or clustering Beatles songs according to their harmonic similarity (Mason, 2012). Others have used the TFIDF scores more directly to identify important melodic fragments in Arab-Andalusian music (Nuttall et al., 2019, 2021).

While we adopt the approach of Nuttall and colleagues in analysing the scores directly, we do not combine TF and IDF in our study for two main reasons. Firstly, there may be more than one route to the same TFIDF score (e.g. low IDF × high TF = low TF × high IDF). Secondly, dynamic expectations are shaped by both the level of repetition and the rarity of a chord progression: a highly repetitive loop implies that dynamic expectations are engaged and confirmed, while a very rare loop implies that dynamic expectations are required. Thus, we keep them separate in the analyses reported below.

2. Materials and methods

2.1. Corpora

Two corpora of mostly American popular music were used in this study. The first was the McGill Billboard Corpus (Burgoyne et al., 2011), which contains 739 unique songs that appeared on the United States Billboard Weekly Hot 100 charts between the years 1958-1991. The songs were transcribed by professional musicians and contain harmonic, formal and metrical information, as well as extensive metadata; however, only the harmonic information was used in this study. Four of the songs contained just one chord throughout; these songs were excluded, giving a final total of 735 songs. The total number of chords was 69,452.

The second was the Yale-Geerdes Billboard corpus, a newly compiled collection of 541 pop songs that appeared on the Billboard Year-End Hot 100 Songs chart between the years 2002 and 2021 inclusive and were available in the catalogue of Geerdes Media (https://geerdes.media/). Geerdes provides highly faithful MIDI transcriptions of commercially available recordings of popular music, capturing performance details (e.g. tempo changes, rhythmic nuances and dynamic changes) as closely as possible to the original audio. Of these 541 songs, 89 came labelled with chord symbols which were extracted using the music21 Python package (Cuthbert & Ariza, 2010). Another 227 songs were furnished with Vocal Harmony (VH) tracks that contained chordal reductions of the musical surface. Chord symbols for these harmonic reductions were extracted using the *chorder* Python package (Chang, 2020). *chorder* uses a simple template-matching procedure to determine the root and quality of a chord up to the seventh. For the remaining 225 songs, each channel of the MIDI file was scanned to find the instrument that contained the highest proportion of chords with three or more notes. If this chord coverage lasted at least 90% of the duration of song, that instrument was selected as the chord track and *chorder* was used once again to extract labels. Four songs from this corpus contained just one chord throughout; therefore, we excluded them, giving a total of 537 songs. The total number of chords was 62,675.

2.2. Chord label pre-processing

The two corpora were encoded with chord quality vocabularies of differing sizes, so it was necessary to re-encode the chords using a common vocabulary of chord qualities. Since more than 91.1% of chords in each corpus were major and minor triads or extensions of these (such as dominant seventh, major ninth and minor seventh chords), we determined that reducing the chord to one of three qualities (major, minor, other) was sufficient.

Subsequently, chord labels were encoded using a tonic-agnostic representation developed in Acevedo (2020). In this encoding scheme, the number of ascending semitones by which the chord root was approached was paired with the chord quality. For instance, a looping passage based on the chord progression F major–C major–G major–A minor (an instance of the Axis progression) would be encoded as $_{8}\text{maj}_{7}\text{maj}_{7}\text{maj}_{2}\text{min}$.

The purpose of using this representation was to account for interval invariance. Two chord progressions with different tonal interpretations might have the same relationship between their roots (e.g., I–V and IV–I), but this similarity would be missed by encoding them with respect to a tonic. Moreover, harmony in popular music is often tonally ambiguous: for instance, the Axis progression may suggest the Aeolian or major mode depending on its rotation, metrical configuration and the melody it accompanies (Richards, 2017).

2.3. Looping passage detection and conversion to loopclasses

For all songs, we detected all looping passages, which we defined as two to eight chords – the loop unit – followed immediately by a full literal repetition and again by at least a partial literal repetition. Subsequently, we converted the loop units to loopclasses, which we defined as a canonical rotation of a loop unit, arbitrarily chosen to be the rotation that is alphabetically first. We recorded the number of iterations of that loopclass in the looping passage and subsequently totalled the number of iterations of that loopclass across all looping passages of a song. An example of the looping passage detection and loopclass conversion procedure is shown in Figure 1. 53.7% of chords in the McGill corpus and 70.5% of chords in the Yale-Geerdes corpus belonged to looping passages.

…maj$_7$maj$_8$maj$_7$maj$_7$maj$_2$min$_8$maj$_7$maj$_7$maj$_2$
min$_8$maj$_2$min$_7$maj$_7$maj$_2$min$_8$maj$_7$maj$_7$maj$_2$min
$_8$maj$_7$maj$_2$min…

⇩

2.5 × **loop unit** $_8$maj$_7$maj$_7$maj$_2$min
2.25 × **loop unit** $_7$maj$_7$maj$_2$min$_8$maj

⇩

5 iterations of **loopclass** $_2$min$_8$maj$_7$maj$_7$maj

Figure 1. Conversion of a looping passage based on the Axis progression to the number of iterations of a loop unit and its associated loopclass.

2.4. Calculating TF and normalised IDF (nIDF)

The TF, which corresponds to the repetitiveness of a loopclass in a song, is calculated using the formula

$$\mathrm{TF}(l, s) = \frac{f_{l,s}}{|s|}$$

where l is a loopclass, s is a song, $f_{l,s}$ is the number of chords in s that are part of looping passages based on l, and $|s|$ is the number of chords in s. The TF can range between 0 and 1.

The IDF, which corresponds to the rarity of the loopclass in the corpus, is calculated using the formula

$$\mathrm{IDF}(l, C) = \log_2 \frac{|C|}{|\{s \in C : l \in s\}|}$$

where l is a loopclass, C is a corpus, $|C|$ is the number of songs in C, and $|\{s \in C\!: l \in s\}|$ is the number of songs in C containing l.

Since the size of the corpora were different, we define the normalised IDF (nIDF) as

$$\mathrm{nIDF}(l, C) = \mathrm{IDF}(l, C) - \max_{l' \in C}\big(\mathrm{IDF}(l', C)\big)$$

where the second term denotes the maximum possible IDF in the corpus. This value is equivalent to $\log_2|C|$, corresponding to a loopclass appearing in just one song. The nIDF is nonpositive and ranges from 0 to $-\log_2|C|$.

2.5. Statistical tests

We compared the TF and nIDF distributions of the two corpora using Mann-Whitney U tests, whose null hypothesis is that the populations under comparison are drawn from the same distribution (without specifying what that distribution is). A rejection of the null hypothesis means that on average one distribution is greater than the other. Where appropriate, we also ran additional tests on subsets of the TF and nIDF distributions, detailed in Sections 3.3 and 3.4.

3. Results

3.1. TF distributions

Figure 2 shows the TF distributions for both corpora. The McGill corpus TF distribution exhibits a unimodal shape, with the densest TF range between 0.1-0.2, accounting for 22.8% of loopclass-song pairs. Meanwhile, the Yale-Geerdes corpus has a bimodal shape, with the densest TF ranges at 0.1-0.2 (17.1% of loopclass-song pairs) and 0.9-1 (21.4% of loopclass-song pairs). The median TFs for the McGill and Yale-Geerdes corpora were .26 and .41 respectively.

The Mann-Whitney test revealed that the Yale-Geerdes distribution was on average greater than the McGill distribution ($U = 372{,}890.5$, $p < .0001$). Thus, songs in the 21st-century corpus are more repetitive on average.

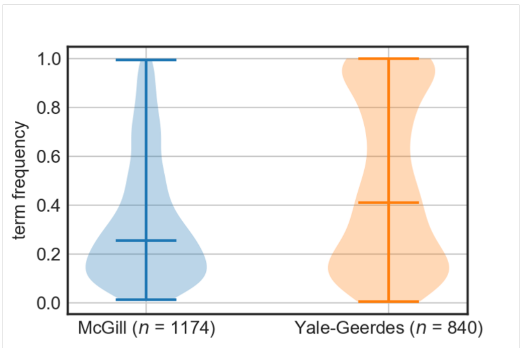

Figure 2. Violin plots of the TF distributions for the McGill and Yale-Geerdes corpora, with medians and the number of loopclass-song pairs indicated.

3.2. nIDF distributions

Figure 3 shows the nIDF distributions for both corpora. Both distributions are concentrated at an nIDF of 0, which was the median IDF for both corpora. In other words, the majority of loopclasses in the corpus appear in only one song (McGill: 79.6%, 367 loopclasses; Yale-Geerdes: 77.0%, 313 loopclasses). The Mann-Whitney test revealed that there was no significant difference between the two nIDF distributions ($U = 96287.5$, $p = .31$).

One other notable aspect of Figure 3 is that the tail of the McGill nIDF distribution stretches much further to the negative than the tail of the Yale-Geerdes nIDF distribution, meaning that the most schematic 20th-century loopclass (a two-chord shuttle between major chords separated by a fifth, $_5$maj$_7$maj) appears in more songs than the most schematic 21st-century loopclass (corresponding to the Axis progression, $_2$min$_8$maj$_7$maj$_7$maj). Therefore, we explored whether

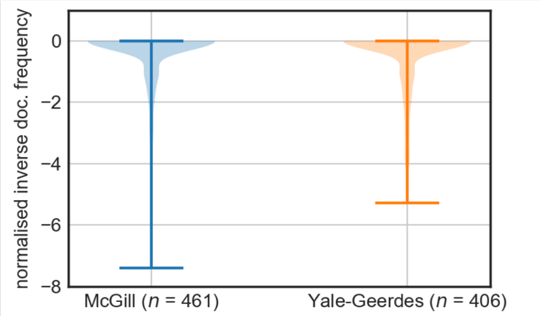

Figure 3. Violin plots of the nIDF distributions for the McGill and Yale-Geerdes corpora, with medians and the number of loopclass-song pairs indicated.

the nIDF distributions for only the schematic loopclasses (i.e. loopclasses appearing in more than one song) were significantly different between corpora. The Mann-Whitney test did not reveal a significant difference ($U = 4021.0$, $p = .33$); indeed, the median schematic nIDF for both corpora was -1.58 (corresponding to a loopclass appearing in three songs).

However, we also conducted a Kolmogorov-Smirnov test, which identifies whether the maximum distance between two cumulative distribution functions (CDFs) is large enough to conclude that the samples are drawn from different distributions. Under this test, the distributions were significantly different ($D = .23$, $p = .01$). The maximum distance between the CDFs was achieved at nIDF = -4.09 (corresponding to a loopclass appearing in 17 songs). For more negative nIDFs, the tail of the McGill distribution was heavier than the tail of the Yale-Geerdes distribution. Thus, the Yale-Geerdes corpus's most common schemata do not appear in as many songs as they do in the McGill corpus.

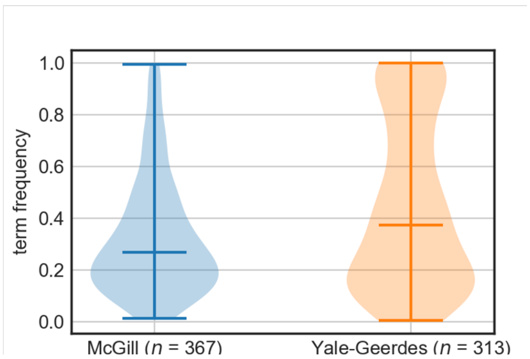

Figure 4. Violin plots of the TF distributions for the one-song McGill and Yale-Geerdes loopclasses, with medians and the number of loopclass-song pairs indicated.

3.3. TF distributions for loopclasses appearing in one song

The nIDF analysis revealed that the majority of loopclasses appear in just one song, meaning that dynamic expectations are required. Thus, we explored how those loopclasses are used in each corpus by examining their TF distributions. These are shown in Figure 4. One-song loopclasses account for 31.2% of loopclass-song pairs in the McGill corpus and 37.3% of loopclass-song pairs in the Yale-Geerdes corpus

The shapes of these distributions are similar to the shapes of the full TF distributions in Figure 2. The median one-song loopclass TF for the McGill corpus was .27 and for the Yale-Geerdes corpus it was .37. The Mann-Whitney test revealed a significant difference between the corpora ($U = 47242.0$, $p < .0001$). Thus, the one-song loopclasses are repeated more in the 21st-century corpus than the 20th-century corpus.

4. Discussion, conclusions and future directions

Three key results emerged from the corpus study reported above. Firstly, loopclasses in the 21st-century Yale-Geerdes corpus are significantly more repetitive than the 20th-century McGill corpus. Secondly, a similarly large majority of loopclasses in each corpus appear in only one song; however, the most schematic loopclasses in the McGill corpus appeared in more songs than the most schematic loopclasses in the Yale-Geerdes corpus. Thirdly, the average repetitiveness of loopclasses appearing in only one song was also significantly higher in the Yale-Geerdes than the McGill corpus.

The increased within-song repetition of loopclasses within the 21st-century corpus suggests the increased importance of dynamic expectations for harmony in recent popular music: on average, they are more readily confirmed than they are in the 20th-century corpus. With 21.4% of Yale-Geerdes loopclass-song pairs having a TF of 0.9-1, the repetition may be sufficient to shift a listener's attention to some other parameter such as melody, which is frequently "divorced" from the chordal background in terms of its pitch content (Covach, 2018; de Clercq & Margulis, 2018; Nobile, 2015; Reed, 2022; Temperley, 2007). There is perceptual evidence that less predictable lines in a musical texture are heard as more prominent than predictable lines (Taher et al., 2016). This is only possible due to the engagement of dynamic expectations in musical structure perception.

While at the song level the Yale-Geerdes corpus was more repetitive, both corpora contained similar proportions of rare and schematic loopclasses as indexed by the IDF distributions. A large majority of loopclasses in each corpus – more than three-quarters in each case – appeared in just one song. However, the most schematic loopclass in McGill corpus appears in more songs than the corresponding loopclass in Yale-

Geerdes corpus. After excluding the loopclasses appearing in only one song, the distributions became significantly different under the Kolmogorov-Smirnov test, owing to the heavier tail of the IDF distribution for the McGill corpus. This suggests that schemata have declined in their informativeness for the formation of expectations in 21st-century popular music compared to 20th-century popular music. In other words, a listener may need to rely on dynamic expectations more often in 21st-century popular music to orient themselves to what they are hearing.

Finally, when we considered just the songs in which dynamic expectations are required – those where the loopclass only occurs in that song – the 21st-century corpus is still more repetitive, with similar TF distributions compared to the whole corpus for both centuries. Thus, when dynamic expectations are absolutely required, they are confirmed more in popular music from the 21st-century than from the 20th-century. In other words, a listener can expect extensive repetition in 21st-century popular music, but it is less clear exactly *what* will be repeated.

It bears noting that while one-song loopclasses are non-schematic by definition, there is a continuum of how schematic a loopclass can be. For instance, suppose a loopclass appears in only two songs. If a listener has heard only one of the two songs in which that loopclass appears, they would still need to engage dynamic expectations as they would not have prior experience with that loopclass. Acevedo (2020) defined a cut-off of 20 songs for an entropy-bounded lick to be classed as schematic, but a making binary distinction is a somewhat arbitrary decision. We draw such a distinction for the purposes of illustrating how one-song loopclasses are used in comparison to the whole corpus (particularly given that more than 30% of loopclass-song pairs in each corpus are one-song loopclasses), but we acknowledge that the corpus is an imperfect proxy for a listening history.

In conclusion, we have provided computational support for the claims in de Clercq & Margulis (2018) that dynamic expectation is important in the perception of popular music harmony, particularly from the 21st century. More broadly, we claim that studies of musical style can be enriched by the consideration of non-normative patterns. In future work, we hope to add nuance to these results by incorporating increased genre and chronological specificity, as well as a measure of similarity between schematic and one-song loopclasses, which is not represented under the current encoding scheme.

References

Acevedo, S. (2020). *Harmonic Schemata of Popular Music: An Empirical Investigation of Analytical Patterns and Their Mental Representations* [Ph.D. Thesis, Yale University].

Aslin, R. N. (2017). Statistical learning: A powerful mechanism that operates by mere exposure. *WIREs Cognitive Science*, *8*(e1373), 1–7.

Bharucha, J. (1987). Music Cognition and Perceptual Facilitation: A Connectionist Framework. *Music Perception*, *5*(1), 1–30.

Biamonte, N. (2010). Triadic Modal and Pentatonic Patterns in Rock Music. *Music Theory Spectrum*, *32*(2), 95–110.

Burgoyne, J. A., Wild, J., & Fujinaga, I. (2011). An Expert Ground Truth Set for Audio Chord Recognition and Music Analysis. *Proceedings of the 12th International Society for Music Information Retrieval Conference*, 633–638.

Chang, J. (2020). *Chorder* (0.1.2) [Python]. https://github.com/joshuachang2311/chorder

Cornelissen, B., Zuidema, W., & Burgoyne, J. A. (2020, October 11). Mode classification and natural units in plainchant. In *Proceedings of the 21st International Society for Music Information Retrieval Conference*. ISMIR.

Covach, J. (2018). Analysing texture in rock music: Stratification, coordination, position and perspective. In R. von Appen & A. Doehring (Eds.), *Pop weiter denken: Neue Anstöße aus Jazz Studies, Philosophie, Musiktheorie und Geschichte*. transcript Verlag.

Cuthbert, M. S., & Ariza, C. (2010). music21: A Toolkit for Computer-Aided Musicology and Symbolic Music Data. *Proceedings of the 11th International Society for Music Information Retrieval Conference*, 637–642.

de Clercq, T., & Margulis, E. H. (2018). A Psychological Perspective on Repetition in Popular Music. In O. Julien & C. Levaux (Eds.), *Over and over: Exploring repetition in popular music* (Paperback edition, pp. 147–161). Bloomsbury Academic.

de Clercq, T., & Temperley, D. (2011). A corpus analysis of rock harmony. *Popular Music*, *30*(1), 47–70.

Doll, C. (2017). *Hearing harmony: Toward a tonal theory for the rock era*. University of Michigan Press.

Duinker, B. (2019). Plateau Loops and Hybrid Tonics in Recent Pop Music. *Music Theory Online*, *25*(4).

Gjerdingen, R. (2007). *Music in the galant style*. Oxford University Press.

Hansen, N. Chr., & Pearce, M. T. (2014). Predictive uncertainty in auditory sequence processing. *Frontiers in Psychology*, *5*.

Huron, D. (2006). *Sweet Anticipation: Music and the Psychology of Expectation*. MIT Press.

Jonaitis, E., & Saffran, J. (2009). Learning Harmony: The Role of Serial Statistics. *Cognitive Science*, *33*(5), 951–968.

Loui, P., Wessel, D. L., & Hudson Kam, C. L. (2010). Humans Rapidly Learn Grammatical Structure in a New Musical Scale. *Music Perception*, *27*(5), 377–388.

Luhn, H. P. (1957). A Statistical Approach to Mechanized Encoding and Searching of Literary

Information. *IBM Journal of Research and Development*, *1*(4), 309–317.
Margulis, E. H., & Beatty, A. P. (2008). Musical Style, Psychoaesthetics, and Prospects for Entropy as an Analytic Tool. *Computer Music Journal*, *32*(4), 64–78.
Mason, D. J. (2012). *The Beatles Genome Project: Cluster Analysis and Visualization of Popular Music*. IEEE VIS, Seattle, WA.
Nobile, D. (2015). Counterpoint in Rock Music: Unpacking the “Melodic-Harmonic Divorce”. *Music Theory Spectrum*, *37*(2), 189–203.
Nobile, D. (2020). *Form as harmony in rock music*. Oxford University Press.
Nuttall, T., Casado, M. G., Ferraro, A., Conklin, D., & Caro Repetto, R. (2021). A computational exploration of melodic patterns in Arab-Andalusian music. *Journal of Mathematics and Music*, *15*(2), 168–180.
Nuttall, T., Casado, M. G., Tarifa, V. N., Repetto, R. C., & Serra, X. (2019). Contributing to new musicological theories with computational methods: The case of centonization in Arab-Andalusian music. *Proceedings of the 20th International Society for Music Information Retrieval Conference*, 223–228.
Pearce, M. (2018). Statistical learning and probabilistic prediction in music cognition: Mechanisms of stylistic enculturation. *Annals of the New York Academy of Sciences*, *1423*(1), 378–395.
Reed, S. A. (2022). An Idiom of Melodic-Harmonic Divorce Sub-Circle Motion in Popular Music. *Music Theory Online*, *28*(2).
Richards, M. (2017). Tonal Ambiguity in Popular Music’s Axis Progressions. *Music Theory Online*, *23*(3).
Sears, D. R. W., & Forrest, D. (2021). Triadic patterns across classical and popular music corpora: Stylistic conventions, or characteristic idioms? *Journal of Mathematics and Music*, *15*(2), 140–153.
Spärck Jones, K. (1972). A Statistical Interpretation of Term Specificity and Its Application in Retrieval. *Journal of Documentation*, *28*(1), 11–21.
Tagg, P. (2014). *Everyday Tonality II: Towards a tonal theory of what most people hear*. The Mass Media Music Scholars’ Press.
Taher, C., Rusch, R., & McAdams, S. (2016). Effects of Repetition on Attention in Two-Part Counterpoint. *Music Perception*, *33*(3), 306–318.
Temperley, D. (2007). The melodic-harmonic ‘divorce’ in rock. *Popular Music*, *26*(2), 323–342.