

## ANTICIPATORY SYNCOPATION IN ROCK: A CORPUS STUDY

IVAN TAN, ETHAN LUSTIG, & DAVID TEMPERLEY  
*Eastman School of Music, University of Rochester*

WHILE SYNCOPATION GENERALLY REFERS TO ANY conflict between surface accents and underlying meter, in rock and other recent popular styles it takes a more specific form in which accented notes occur just before strong beats. Such “anticipatory” syncopations suggest that there is an underlying cognitive representation in which the accented notes and strong beats align. Syllabic stress is crucial to the identification of such syncopations; to facilitate this, we present a corpus of rock melodies annotated with lyrics and syllabic stress values. We propose a new measure of syncopation that incorporates syllabic stress; we also propose a measure of anticipatory syncopation, and show that it reveals a strong presence of this type of syncopation in rock music. We then use these measures to explore other aspects of syncopation in rock, including its occurrence in different parts of the 4/4 measure, its dependence on tempo, its historical evolution, and its aesthetic functions.

*Received: January 30, 2018, accepted November 20, 2018.*

**Key words:** rhythm and timing, popular music, metre, corpus, music and language

SYNCOPATION, AS THE TERM IS GENERALLY understood, refers to a conflict between the accents of a piece and the underlying meter. In Figure 1A, the notes marked with asterisks in the second and fourth measures can be regarded as syncopations. The mere fact that there are note onsets on weak (8th-note) beats, in contrast to the stronger beats in between (at the quarter-note level and higher), confers a kind of accent on those weak beats. Thus, the rhythm could be said to conflict with the meter—under the assumption that it is most normative for strong beats to be more accented than weaker ones. The fact that the starred notes are not followed by any note on the following quarter-note beats also increases their length, which accents them in another way (sometimes known as *agogic* or *durational* accent). (Here, and throughout the study, we follow the common practice of defining the “length” of a note as its *interonset*

*interval* or IOI—the rhythmic interval between the note’s onset and the onset of the following note; thus a rest is absorbed into the previous note.) Defined in this way, syncopation is a common phenomenon in classical music and many other styles, widely recognized in music theory (Krebs, 1999; Lerdahl & Jackendoff, 1983; London, 2012).<sup>1</sup>

In rock and other kinds of recent popular music, syncopation takes on a rather different character, as illustrated by Figure 2A. Here, as in Figure 1A, certain notes (marked again with asterisks) fall on relatively weak beats, with no notes on the following strong beats. In addition, however, there is a sense that these notes *anticipate* the following strong beats—that they belong on these beats in some way, or at least are associated with them. This is made clear if we shift the syncopated notes one 8th-note to the right, as shown in Figure 2B; now these notes fall on stronger beats. Of course, similar shifts could be applied to Figure 1A as well, with similar results (see Figure 1B). But there are other motivations for shifting the notes in Figure 2 that do not apply to Figure 1. In Figure 2A, the linguistically stressed syllable “son” and the unstressed syllable “my” both fall on weak 8th-note beats; but once these syllables are shifted, “son” falls on a stronger beat than “my.” Syllabic stress is a well-known source of musical accent; in many styles (such as classical music and European folk music), stressed syllables of text usually fall on strong beats (Halle & Lerdahl, 1993; Temperley & Temperley, 2013). Shifting the syllables of Figure 2A (see Figure 2B) brings the melody into accordance with this general principle. By contrast, shifting the syllables in Figure 1A does not have this effect (see Figure 1B); indeed, the unstressed syllable “and” in the second and fourth measures now falls on a stronger beat than the surrounding stressed syllables, worsening the alignment between meter and stress rather than improving it. Yet another motivation for shifting the notes of Figure 2 is the fact that the final note, C#, fits better with the F# minor harmony of the final measure than with the previous

<sup>1</sup> The reader may wonder why we chose an example from 20th-century art music, rather than from common-practice music. The reason for this is that in music of the common-practice period, syncopations in *vocal* music appear to be extremely rare; we were unable to find a good English-language example.

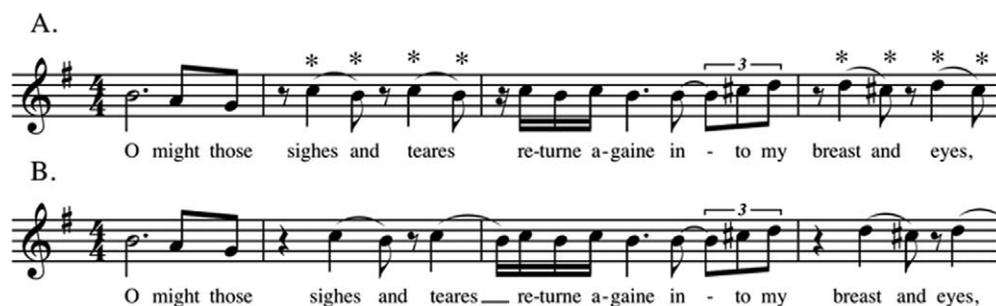


FIGURE 1. Britten, “O might those sighes and teares,” from *The Holy Sonnets of John Donne*, Op. 35. (A) Mm. 3-6; (B) recomposed with syncopation removed.

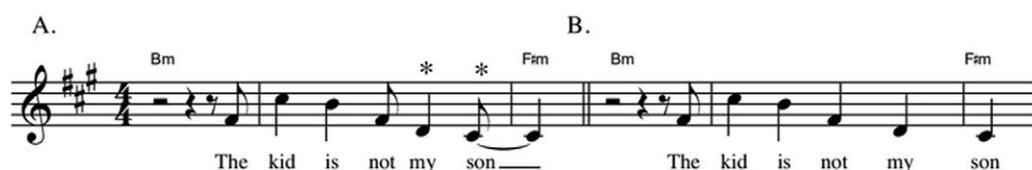


FIGURE 2. Michael Jackson, “Billie Jean,” last line of chorus. (A) Original rhythm; (B) recomposed with syncopation removed.

B minor harmony (it is a chord-tone of F# minor but not of B minor); thus, shifting it to the following downbeat makes harmonic sense. Again, this reason for regarding the syncopated notes in Figure 2 as anticipating the following strong beats is not present in Figure 1 (though this is perhaps debatable; the harmony expressed by the accompaniment—not shown here—is rather ambiguous).

The kind of syncopation observed in Figure 2—what we will call *anticipatory syncopation*—has been noted and discussed by a number of scholars as an important feature of rock and other 20th-century popular styles, such as ragtime, jazz, and the blues (Fox, 2002; Kono-witz, 1991; Temperley, 1999; Tilton, 1994). It has rarely been the subject of focused attention; this is reflected in the fact that there is no widely accepted term for it.<sup>2</sup> Moreover, the validity of this concept is not universally accepted. Many discussions of rock and related styles simply describe syncopation as a conflict between accents and meter—thus adopting the more general understanding of the term, without any notion of anticipation. Everett (2009), describing syncopation in rock, describes it simply as a “clash of foreground against background”; Stephenson (2002) offers a similar view.

<sup>2</sup> A search on Google Scholar suggests that the phrase “anticipatory syncopation” has only been used occasionally and in passing (to describe specific pieces), never in general discussions of rhythm or musical styles. The term “forward syncopation” is also occasionally used (e.g., Jenness & Velsey, 2014).

Biamonte (2014) characterizes syncopation in rock as a kind of “rhythmic dissonance”—again indicating a conflict between surface rhythm and meter, but without any implication that anticipation is involved. As another example, in the *New Grove Dictionary of Jazz* (Kernfeld, 2002), the article on “beat” defines syncopation as “the shifting of articulations from stronger beats to weaker ones or to metrical positions that do not fall on any of the main beats of the bar”; there is no suggestion that syncopated notes tend to anticipate the following strong beats.

A fundamental question arises here about the mental representation of rhythm in rock (and other popular styles). If we view syncopations like those in Figure 2A as “belonging” on the following beats, this suggests that there is some kind of underlying unsyncopated cognitive representation (like that shown in Figure 2B) involved in the production and perception of such rhythms, and that surface rhythms are derived by shifting the onsets in relation to this underlying representation. The psychological reality of such a claim is not obvious, and cannot be demonstrated simply by a few examples. Syncopations like those in Figure 2A might sometimes occur just by chance, even if the musicians were not thinking of them as anticipatory. In this article, we aim to determine the extent to which anticipatory syncopation (and, by extension, the kind of underlying representation posited in Figure 2B) is operative in the mental representation of rock rhythm. We first propose a new measure of syncopation that incorporates syllabic



FIGURE 3. The Police, “Canary in a Coalmine,” beginning of first verse.

stress; we then propose a second measure that specifically quantifies anticipatory syncopation. To facilitate the application of these measures, we have created a corpus of rock melodies annotated with lyrics and stress values, which is publicly available and may be useful for other purposes as well. We use our measures to quantify the amount of syncopation in our rock corpus (comparing it to a small corpus of 19th-century English songs) and also the amount of anticipatory syncopation. Finally, we consider some other issues to which our quantitative measures of syncopation and anticipatory syncopation might be applied.

A number of proposals have been offered for how to quantify syncopation (for a review, see Gómez, Thul, & Toussaint, 2007); related to this is the idea of rhythmic complexity, which is sometimes taken to be more or less synonymous with syncopation (Pressing, 1999; Smith & Honing, 2006). These previous proposals all relate to the general concept of syncopation, rather than specifically anticipatory syncopation, so they do not require detailed discussion here. Most of these models define syncopation purely in terms of the locations of events in relation to the metrical structure, not considering other sources of accent; we might call these “positional” models of syncopation. Perhaps the most well-known such model is that of Longuet-Higgins and Lee (1984), which defines syncopation as a note on a beat followed by a rest (or a continuation of the note) on the next beat, with the second beat stronger than the first. The “strength” of the syncopation increases as the metrical strength of the syncopated note’s beat decreases, and increases with the strength of the following rest; that is, a syncopation is stronger if it falls on a weak beat and precedes a much stronger beat. Similarly, Huron and Ommen (2006) define syncopation as an event on a weak beat with a rest (or continuation) on the following strong beat. They present corpus evidence that syncopation, as they define it, increased in frequency from the 1890s to the 1930s in American popular music.

Both the Longuet-Higgins/Lee (1984) and Huron/Ommen (2006) models define a syncopation, essentially, as a weak-beat note followed by a strong-beat rest or continuation. This definition embodies the traditional idea of syncopation as a conflict between

accent and meter in two ways: because there is a note on a weak beat and not on a neighboring strong one, and because the note on the weak beat is relatively long. While this purely positional conception of syncopation is valuable—indeed, our own model incorporates it—it neglects other important aspects of syncopation. In Figure 3, the last syllable of the first phrase (marked with an asterisk) would be regarded as a syncopation by the abovementioned models—indeed, by Longuet-Higgins and Lee’s model, it is a very strong syncopation, since the note’s beat is weak and the following beat is very strong (a downbeat). Yet, in our view, there is little if any sense of syncopation here, anticipatory or otherwise. Of crucial importance here is syllabic stress. The syllable “-fect” is unstressed in relation to the previous syllable “per-,” and thus it is appropriate for “-fect” to be on a weaker beat. It is true that this syllable is relatively long (in relation to surrounding notes), which arguably confers a slight agogic accent on it. But this is outweighed by the weak stress level of the syllable. It is because of situations like this that we feel it is crucial for syllabic stress to be incorporated into any satisfactory measure of syncopation in vocal melody.

Two recent studies are important precedents for the current project: those of Condit-Schultz (2016) and Waller (2016). Both of these authors explore rhythm in hip-hop, taking a probabilistic, corpus-based approach, and both authors incorporate syllabic stress. Waller notes the close connection between syncopation and complexity; he uses entropy to quantify the complexity of rap rhythms, including syllabic stress as a factor. Condit-Schultz employs a measure of syncopation similar to that of Longuet-Higgins and Lee, but applies it to the stressed syllable rhythmic layer only, ignoring all unstressed syllables. Our approach builds on these earlier studies in incorporating syllabic stress as an aspect of melodic rhythm; however, neither Waller nor Condit-Schultz attempts to distinguish anticipatory syncopation from syncopation more generally.

### The Corpus

In this study we use the Rolling Stone corpus (modified in several ways, as described below) as the dataset for our

A.

Hey Jude Don't make it bad Take a sad song And make it bet-ter Re-

B.

[F] [OCT=4] ...5 | 3....356 | v2.....23 | 4.^1..175 | 6.543.....5. |  
 with melisma indicated: | 6.(543).....5. |

FIGURE 4. The Beatles, “Hey Jude,” beginning. (A) In traditional notation; (B) as coded in symbolic notation in the RS corpus.

analyses. (The various RS corpus files discussed in this paper, including the new stress-annotated corpus presented here, are all available at [rockcorpus.midside.com](http://rockcorpus.midside.com).) The Rolling Stone corpus (hereafter the RS corpus) includes harmonic analyses and melodic transcriptions of 200 songs, a subset of *Rolling Stone* magazine’s list of the “500 Greatest Songs of all Time” (de Clercq & Temperley, 2011; Temperley & de Clercq, 2013). The corpus offers a diverse sample of “rock” songs, broadly construed, from the 1950s through the 1990s. Several other corpora of popular music have also been created (Bertin-Mahieux, Ellis, Whitman, & Lamere, 2011; Burgoyne, Wild, & Fujinaga, 2011; Mauch et al., 2009); while the RS corpus is smaller than these other corpora, it is the only one that includes transcriptions of vocal melodies.

In the vocal melodies of the RS corpus, notes are represented as scale degrees (pitch classes in relation to the tonic). Figure 4A shows the beginning of the Beatles’ “Hey Jude” in traditional notation; in the RS corpus, this excerpt is represented as shown in the symbolic notation in the top line of Figure 4B. Each note is assumed to be the closest registral instance of that scale degree to the previous note, unless marked with *v* (down an octave) or *^* (up an octave). (Tritone intervals are assumed to be ascending unless marked with *v*.) Vertical bars represent barlines. Each measure is divided equally into a number of rhythmic units (usually 4, 8, or 16); a dot indicates a unit with no note onset. [F] indicates the key, while [OCT=4] indicates the octave of the first note (following the usual convention, where middle C is the lowest note of octave 4). We see that the melody begins on scale degree 5. From these three pieces of information (key, octave, scale degree) we know that the first pitch is C4, and from there, we can derive all the remaining pitches of the melody. While traditional notation distinguishes between sustained notes and rests (requiring the transcriber to make decisions about note duration), the RS corpus only encodes note onsets, not offsets. As discussed

by Temperley and de Clercq (2013), there is some subjectivity with regard to the transcription of both pitch and rhythm, the latter being especially relevant in this context; for example, it is sometimes a judgment call whether a note’s deviation from a strong beat is a true syncopation, or simply an expressive nuance of “micro-timing.”

For the current study, we expanded a subset of the RS corpus by adding melismas, lyrics, and syllabic stress data. We started with a 100-song subset of the corpus, taking the 20 highest-ranked songs from each decade from the 1950s through the 1990s to give us a balanced representation of decades. From these, we excluded 20 songs: 17 songs that featured time signatures other than 4/4 or 2/4, and three songs without sung melody (rap songs). The remaining 80 songs constitute the corpus used in this study. We then marked melismas (defined as a continuously sung syllable involving two or more pitches) in the melodies, which we indicated by placing the notes of a melisma inside parentheses. For instance, in Figure 4A, “-ter” in “better” spans three pitches; this melisma is represented as shown on the second line of Figure 4B. Marking melismas greatly facilitated the alignment of the melodies with lyrics; once non-initial melisma notes are excluded, each syllable may be mapped onto a single note.

Having marked melismas, we then downloaded the lyrics of each song from the Internet. A number of websites provide large libraries of song lyrics; after considering several options, we chose [chartlyrics.com](http://chartlyrics.com), since it had the most complete lyrics for all of the songs in the corpus. Where necessary, the lyrics were edited for correctness and consistency, and other lyrics sources consulted for corroboration. Finally, we used the Carnegie Mellon University (CMU) Pronouncing Dictionary ([speech.cs.cmu.edu/cgi-bin/cmudict](http://speech.cs.cmu.edu/cgi-bin/cmudict)) to assign a stress value to each syllable. The dictionary accepts English words as input, and outputs one of three possible stress values to each syllable in each word. Unstressed syllables receive a 0; primary stress receives a 1; secondary stress

1	2	3	4	5	6
hey_jude		404	0.75	60	1 HEY[1]
hey_jude		404	1	57	1 JUDE[1]
hey_jude		404	1.625	57	1 DONT[1]
hey_jude		404	1.75	60	1 MAKE[1]
hey_jude		404	1.875	62	0 IT[1]
hey_jude		404	2	55	1 BAD[1]
hey_jude		404	2.75	55	1 TAKE[1]
hey_jude		404	2.875	57	0 A[1]
hey_jude		404	3	58	1 SAD[1]
hey_jude		404	3.25	65	1 SONG[1]
hey_jude		404	3.625	65	0 AND[1]
hey_jude		404	3.75	64	1 MAKE[1]
hey_jude		404	3.875	60	0 IT[1]
hey_jude		404	4	62	1 BETTER[1]
hey_jude		404	4.125	60	0 BETTER[2]
hey_jude		404	4.875	60	0 REMEMBER[1]

FIGURE 5. Final output representation of the expanded RS corpus.

receives a 2, and only occurs in words that also contain a primary stress. For example, the word “music” is 10, while “dictionary” is 1020.

In some cases, a single orthographic form represents two different words with different stress patterns, e.g., “I broke a RE-cord” and “I will re-CORD the music.” In such cases, the CMU Dictionary contains two different entries, and we hand-encoded the stress patterns accordingly. We also added some words to the dictionary, such as uncommon words and names (“hedge-row,” “Maybellene”), and variant pronunciations (e.g., “tryin” vs. “trying”—or even the monosyllabic “tryn”). We created a separate category for non-words (like “uh,” “ah,” and “oh”), giving them a stress value of 3.

The CMU Dictionary assigns all monosyllabic words a stress value of 1, with the sole exception of “and,” “a,” and “the,” which receive a 0. It is generally agreed, however, that in English, monosyllabic “function words”—such as prepositions, conjunctions, and auxiliary verbs—are normally unstressed (Hayes, 1995; Kelly & Bock, 1988; Selkirk, 1996). To remedy this issue, we assigned a value of 0 to all monosyllabic function words, as defined by a pre-existing word list (see [sequencepublishing.com/1/academic.html](http://sequencepublishing.com/1/academic.html)). This list includes conjunctions (e.g., “and,” “but”), determiners (e.g., “a,” “the”), prepositions (e.g., “at,” “on,” “to”), pronouns (e.g., “me,” “it,” “my”), quantifiers (e.g., “none,” “some”), and some auxiliary verbs (e.g., “can,” “may,” “will”—but not forms of “be,” “have,” and “do,” which only sometimes function as auxiliaries). Based on this list,

31% of the word tokens in the corpus were identified as monosyllabic function words. As an example, the original encoding in the CMU Dictionary would represent “Give it to me” as 1 1 1 1; given our modification, “Give it to me” now yields 1 0 0 0. We also added to the list certain monosyllabic contractions that seem to be generally unstressed, such as “I’m” and “we’ll.”

There are many complexities in English stress patterning. Sometimes, even the same word can be realized differently depending on the context: e.g., “I’m on page thir-TEEN,” “I saw THIR-teen men” (Hayes, 1995). Function words—normally unstressed—may be stressed in certain contexts; for example, the phrase “give it to me,” mentioned earlier, might well be spoken with a stress on “me” (1 0 0 1). In “Billie Jean,” in the line “she’s just a girl who claims that I am the one” (which occurs immediately before the melody shown in Figure 2), one might argue that the pronoun “I” is stressed. We did not attempt to encode these distinctions in our corpus, since we lacked a principled, consistent way of doing so. An alternative approach would be to label the stress of each syllable manually (see Condit-Schultz, 2016, for an example of this approach); while this method has advantages, it introduces the risk of experimenter bias, and also the danger that the judgment of a word’s stress might be affected by its musical context.

Figure 5 shows the final output representation for the corpus, demonstrated using the excerpt in Figure 4B. Each syllable/note event receives its own row. The six numbered columns show:

1. Song title
2. Time signature (4/4 is coded as 404, 2/4 is coded as 204)
3. Within-song position (in the form X.Y, where X = the measure number (starting from 0) and Y = the proportional position within the measure. For instance, beat 4 of measure 1 is 0.75; beat 3.5 of measure 2 is 1.625)
4. Pitch (in MIDI number, where 60 = C4)
5. 4-level stress assignment (0: unstressed, 1: primary stress, 2: secondary stress, 3: non-word)
6. Word (the number in square brackets indicates the syllable number within the word)

Linguistic stress is generally assumed to be hierarchical, extending above words to larger units (Hayes, 1995). In a phrase or sentence, the most stressed syllable is typically near the end; for example, in the line “Take a sad song and make it better” from “Hey Jude” (Figure 4), the primary stress is on “bet-.” Such distinctions between stressed syllables are not encoded in our corpus (unless they are within the same word). As has been noted previously (Temperley, 1999), the connection between linguistic stress and musical meter is primarily local; there is little tendency for higher-level stress distinctions to be reflected in metrical placement. (In “Hey Jude,” “bet-” is on a hypermetrically weak downbeat and thus less metrically accented than the previous downbeat “sad,” but there is little sense of conflict between meter and stress.) For the most part, then, the low-level stress distinctions encoded in our corpus are sufficient to capture intuitions about alignment or conflict between stress and meter. Distinctions between stressed syllables are sometimes important, however; in the following discussion, we rely on our intuitions as to when such distinctions are warranted, though we realize that this is somewhat subjective.

### Measuring Syncopation and Anticipatory Syncopation

The 80 songs of the stress-annotated RS corpus (represented using the format in Figure 5) contain a total of 23,129 notes. (Since non-initial melisma notes are excluded, syllables and notes are in a one-to-one correspondence.) In the following analyses, notes in 2/4 measures—less than 0.5% of the total—were excluded due to the difficulty of defining their metrical positions; thus we consider only notes in 4/4 measures, a total of 23,032 notes. One and a half percent (1.5%) of the notes in 4/4 measures did not fall on any 16th-note beat (these were mostly triplet divisions); we include these in the total

number of notes, but say nothing more about them. (We should note that the assignment of metrical levels within a song—e.g., deciding which level is the quarter-note level—is sometimes debatable; we say more about this below.) With regard to stress, both 1’s (primary stress, 43.5% of all syllables) and 2’s (secondary stress, 1.8%) are treated as stressed; 0’s (unstressed, 50.3%) and 3’s (non-words, 4.4%) are treated as unstressed.

Our aims in this section are twofold: first, to develop a quantitative measure of syncopation, and second, to develop a measure of anticipatory syncopation. Both of our measures combine the traditional “positional” approach to syncopation, described earlier, with syllabic stress. While the quantification of syncopation has been considered before, the quantification of anticipatory syncopation has not. Our hypothesis is that rhythmic patterns in rock melodies are generated from underlying, unsyncopated rhythmic representations, and that certain notes are then shifted by one beat to precede their underlying positions; in common-practice music, by contrast, this anticipatory shifting does not occur. If anticipatory syncopation does occur in rock, how would we expect to see it reflected in corpus data?

As a starting point, we can examine the distribution of note onsets on 16th-note positions of the measure. In common-practice music, melodic onset distributions tend to strongly reflect the metrical structure: the stronger the metrical position, the more notes occur at that point (Huron, 2006; Palmer & Krumhansl, 1990). In Figure 6A, the dotted line shows the onset distribution for a small corpus of 19th-century English songs.<sup>3</sup> The correspondence between metrical strength and onset frequency is clearly evident: the most onsets fall on the downbeat (position 1), followed by the third quarter-note beat (position 9), then the second and fourth quarter-note beats (positions 5 and 13), then the weak 8th-note beats (positions 3, 7, 11, and 15), and finally the weak (even-numbered) 16th-note positions. If rock features the same sort of underlying rhythmic representations as common-practice music, but with a high degree of syncopation, we might expect to find a higher incidence of notes on weaker beats than in common-practice music, since some notes on stronger beats in the underlying representation should be shifted to

<sup>3</sup> The 19th-century English song corpus, created by Temperley and Temperley (2013), contains 10 songs (1397 notes) from the book *Songs of England* (Vol. 1). We use it here because, like the RS corpus, it is annotated with stress data, and non-initial melisma notes are excluded. The stress annotations are based on a function-word list very similar to the one used with the RS corpus, though not identical; in particular, forms of the verb “be” were treated as unstressed in the 19th-century corpus, but stressed in the RS corpus.

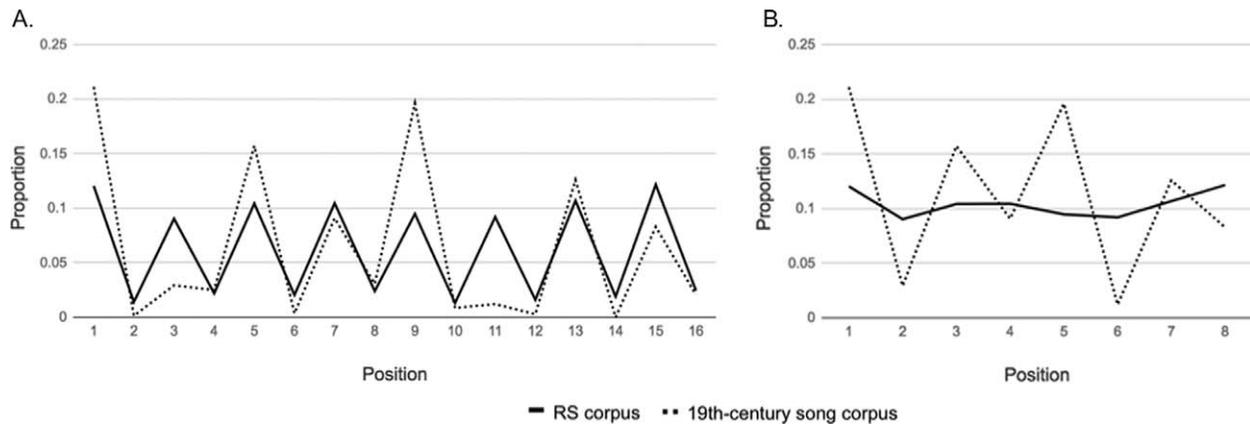


FIGURE 6. (A) Proportion of total syllables at each 16th-note position in the RS and 19th-century corpora. (B) Proportion of total syllables at each 8th-note position in the RS and 19th-century corpora.



FIGURE 7. Chuck Berry, "Roll Over Beethoven," beginning of first verse.

weaker beats in the surface structure. The solid line in Figure 6A shows the frequency distribution of onsets in the RS corpus. As in the 19th-century corpus, there are far more notes on the odd positions (i.e., the 8th-note positions) than on the even positions (the weak 16th-note positions). Among the 8th-note positions, however, the distribution is much flatter than that of the 19th-century song corpus, showing little evidence of any kind of metric differentiation. Figure 6B makes this point clearer by representing only the 8th-note positions.

While the rather flat distribution of onsets (across 8th-note positions) found in the RS corpus could be taken as evidence of syncopation, we should be cautious about drawing this conclusion. An alternative explanation is that the distribution simply reflects dense patterns of 8th-notes. This is sometimes seen in common-practice melodies, and in rock melodies as well; an example is shown in Figure 7. In itself, then, this pattern does not provide clear evidence of syncopation.<sup>4</sup>

<sup>4</sup> It can be seen in Figure 6B that the 19th-century song distribution reflects a higher number of onsets at (8th-note) positions 4 and 8 than at positions 2 and 6. One might wonder if this represents anticipatory syncopation, but we believe that in this case it does not. Rather, we think it reflects the tendency for notes on positions 1 and 5 to be *long*, since they are metrically strong; this means that the following note is more likely to be on the fourth 8th of the half-measure than on the second 8th. As will be seen below, closer study suggests that there is little if any anticipatory

As noted earlier, patterns of syncopation are brought out more clearly when distinctions of syllabic stress are considered. Let us consider the distribution of onsets in the RS and 19th-century corpora, but limited to stressed syllables only. In the case of the RS corpus, we predicted that this distribution would show stronger evidence of the conventional metrical hierarchy than the distribution over all syllables, because stressed syllables presumably adhere to this hierarchy more than unstressed syllables. If a high level of syncopation is present in the RS corpus, however, the alignment should not be as strong as in the 19th-century corpus. We also predicted that this approach would show evidence of anticipatory syncopation in the RS corpus; if there are more stressed syllables on 8th-note positions 1 and 5 (the strong quarter-note beats) than positions 3 and 7 (the weak quarter-note beats) in the underlying cognitive representation, then anticipatory syncopation should yield more stressed syllables on positions 4 and 8 (which precede positions 5 and 1) than on positions 2 and 6 (which precede positions 3 and 7).

syncopation in the 19th-century corpus. One might also wonder if the same phenomenon (the tendency for notes on 1 and 5 to be long) might partly explain the higher frequency of notes at positions 4 and 8 than on 2 and 6 in the rock corpus; but this explanation does not account for the fact that this tendency is much stronger for stressed syllables than unstressed ones, as we will show below.

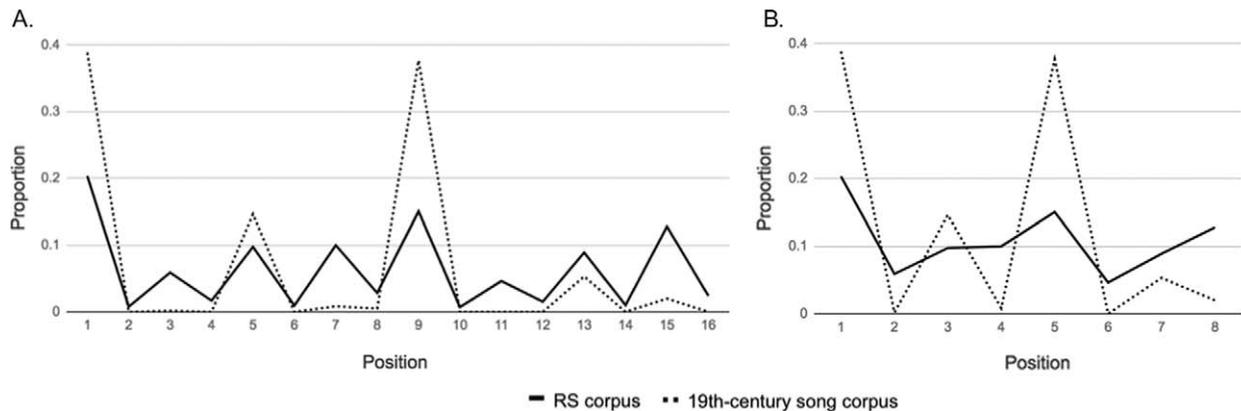


FIGURE 8. (A) Proportion of stressed syllables at each 16th-note position in the RS and 19th-century corpora. (B) Proportion of stressed syllables at each 8th-note position in the RS and 19th-century corpora.



FIGURE 9. (A) The Beatles, "Yesterday," final line of second verse. (B) The Eagles, "Hotel California," beginning of first verse.

Figure 8 shows the distribution of stressed syllables in both corpora, for 16th-note positions (A) and then for 8th-note positions (B). The onset distribution of stressed syllables in the RS corpus does indeed reflect the metrical grid more strongly than the overall onset distribution shown in Figure 6. Like the 19th-century song corpus, the 8th-note positions with the most onsets in the RS corpus are now 1 and 5 (see Figure 8B). However, there is a much higher proportion of stressed syllables on weak 8th-note beats in the RS corpus than in the 19th-century corpus, a highly significant difference, Pearson's  $\chi^2(1) = 240.1$ ,  $p < .0001$ ; this in itself seems to indicate a higher level of syncopation in the RS corpus. In addition, we see in the RS corpus a higher proportion of stressed onsets on 8th-note positions 4 and 8 (which precede the strong quarter-note beats) than on positions 2 and 6 (which precede the weak quarter-note beats). (A way of quantifying this phenomenon, and assessing its statistical significance, will be presented below.) As suggested earlier, this pattern seems to point especially to *anticipatory* syncopation: if syncopation simply involved accenting any weak 8th-note position with equal probability, it is difficult to see why there would be more stressed syllables on positions 4 and 8 than on positions 2 and 6. In the 19th-century song corpus, by contrast, there is little evidence of differentiation between weak 8th-note positions.

(There may be a slight tendency for more onsets at positions 4 and 8 than at positions 2 and 6, but the numbers are too small to draw firm conclusions about this; there are only 18 stressed syllables on weak 8th-note beats in the entire corpus.)

One might consider defining a syncopation as a stressed syllable on a weak 8th-note beat (positions 2, 4, 6, or 8). This would be an oversimplification, however. In Figure 9A, the word "came" is on a weak 8th-note beat and is a stressed syllable (a monosyllabic verb), but it feels less linguistically stressed than the following syllable "SUDD-enly," so it is appropriate for "came" to be on a weaker beat. (Such distinctions between stressed syllables in different words are not encoded in the RS corpus, and they are admittedly somewhat subjective.) Thus, this is not evidence of syncopation. In Figure 9B, "de-" is a stressed syllable on a weak 8th-note beat, but it is on a stronger beat than the following unstressed syllable on a weak 16th-note beat ("-sert"), and feels less stressed than the syllable on the following quarter-note beat ("high-"), so again, its metrical placement is appropriate; thus, there is no reason to posit syncopation. Ideally, a definition of syncopation would not include such cases.

It can be seen that in both of the cases in Figure 9, the starred weak-beat syllable is immediately followed by another syllable on (or even before) the next strong beat. Counting all stressed syllables on weak 8th-note

positions as syncopations adds many false positives of this kind. To exclude such cases, we could limit our search to syllables that are not closely followed in this way. Here we incorporate the positional approach to syncopation, advanced by Longuet-Higgins and Lee (1984) and Huron and Ommen (2006); as discussed earlier, those authors define a syncopation as a weak-beat note with a rest (or continuation of the note) on the following strong beat. We adopt this idea here, but with the added requirement that the syllable must be stressed. While we focus here on the 8th-note level, the definition can also be extended to other metrical levels. Thus, we propose the following operational definition of a syncopation for our study:

*Definition: A syncopation at the  $n$ th-note level is a stressed syllable on a weak  $n$ th-note beat that is not followed by another syllable on or before the next  $n$ th-note beat.*

A further problem is how to evaluate the overall amount of syncopation in a given song or corpus. Simply counting the number of syncopations is unsatisfactory, because a longer melody will tend to have more syncopations; the count must be normalized in some way. There are various ways this could be done, but a solution that yields intuitively good results is simply to divide the number of syncopations by the total number of stressed syllables. We call this the syncopation quotient, or SQ (assuming the definition of syncopation presented above):

$n$ th-level SQ = (total number of syncopations on weak  $n$ th-note beats) / (total number of stressed syllables)

For the 8th-note level specifically, where  $S_8(X)$  = the number of syncopations on 8th-note position  $X$  of the measure:

$$8\text{th-level SQ} = (S_8(2) + S_8(4) + S_8(6) + S_8(8)) / (\text{total number of stressed syllables})$$

Our corpus has 2,382 syncopations at the 8th-note level, out of a total of 10,433 stressed syllables, for an 8th-level SQ of .228. By comparison, the 19th-century corpus contains zero syncopations out of 602 stressed syllables, yielding an 8th-level SQ of 0. This suggests, as expected, that rock features a much higher degree of syncopation than common-practice music (at least, as represented by 19th-century English song).

The definition above concerns syncopation generally, not specifically anticipatory syncopation. For example, several of the syllables in the Britten excerpt presented in

Figure 1—“sighes,” “teares,” “breast,” and “eyes”—are syncopations according to our definition, but (as argued earlier) they do not appear to be anticipatory. To capture the degree to which the syncopation in a corpus is anticipatory, we invoke an observation made earlier: the use of anticipatory syncopation should result in a higher frequency of syncopations on 8th-note positions 4 and 8 than on positions 2 and 6. We can capture this by examining the number of syncopations on positions 4 and 8 as a proportion of all weak-beat syncopations. Thus we propose the anticipatory syncopation quotient, or ASQ, defined as follows:

$$8\text{th-level ASQ} = (S_8(4) + S_8(8)) / (S_8(2) + S_8(4) + S_8(6) + S_8(8))$$

Note that the denominator of the ASQ is the numerator of the SQ. An 8th-level ASQ of more than .5 means that more than half of the 8th-level syncopations are on positions 4 and 8, suggesting that some degree of anticipatory syncopation is present. The 8th-level ASQ for the RS corpus taken as a whole is .772. (We cannot compute an ASQ for the 19th-century corpus, since there are no syncopations at all; the denominator would be zero.) We examined the ASQs for individual songs in the RS corpus (excluding four cases where the denominator is zero), counting the number of ASQs above and below .5 (four cases where the ASQ was exactly .5 were divided between the two categories); 66 out of the 76 ASQ values were above .5, significantly more than half,  $\chi^2(1) = 41.3$ ,  $p < .00001$ .

Notice that we do not propose an operational definition of “anticipatory syncopation.” As observed earlier, it is difficult to say with certainty whether any particular syncopation is truly anticipatory, in the sense that it is understood (by creators and listeners of the song) as belonging on (shifted from) the following strong beat. But if a corpus reflects a much higher incidence of stressed syllables on positions 4 and 8 than on positions 2 and 6, this suggests to us that many of these syllables are anticipatory syncopations, since we can see no other plausible explanation for this pattern.

In at least one respect, our definition of syncopation is too narrow. In some syncopations, the unstressed syllable after the shifted stressed syllable is also shifted; Figure 10A shows a case in point. Since the stressed syllable “fun-” is followed by another syllable (“-ny”) on the next strong beat, it would not be considered a syncopation by our definition. To us, this seems like a clear case of anticipatory syncopation: both syllables of “funny” are shifted to precede the 8th-note beats on which they belong. Indeed, this seems like a rather *extreme* form of

A. I laughed at love 'Cause I thought it was fun-ny

B. Got - ta have some hot stuff Got -

C. Woke up Fell out of bed

FIGURE 10. (A) Jerry Lee Lewis, “Great Balls of Fire,” beginning of second verse. (B) Donna Summer, “Hot Stuff,” first chorus. (C) The Beatles, “A Day in the Life,” beginning of bridge section.

A. And a-ny - time you feel the pain

B. And a-ny - time you feel the pain

FIGURE 11. The Beatles, “Hey Jude,” beginning of bridge. (A) Original rhythm; (B) recomposed with syncopation removed.

syncopation, giving the phrase a more syncopated feel than shifts involving a stressed syllable alone. (Perhaps the reason for this effect is that the conflict between stress and meter is especially acute: the unstressed syllable is on a strong quarter-note beat, and the stressed syllable is on a weak 8th-note one.) Notably, this pattern would not be considered a syncopation at all by the positional criterion. (It *would* be considered a syncopation according to Condit-Schultz’s [2016] definition, which ignores all unstressed syllables.) In our corpus, however, such patterns are relatively rare, and identifying them as syncopations leads to a large number of false positives (often due to the complexities in labeling stresses discussed earlier) so we do not include them in the current measure. However, this phenomenon deserves further study; it suggests to us that conflicts between syllabic stress and meter may be an even greater factor in the perceived “strength” of a syncopation than the positional factors discussed by Longuet-Higgins and Lee and others.

Our definition of syncopation also results in some false positives; an example is shown in Figure 10B. “Stuff” is a stressed syllable on a weak 8th-note beat, so it would be considered a syncopation by our definition, but it is less stressed than the previous syllable “hot” and on a weaker beat. Since there is no conflict between stress and meter, one of the main motivations for regarding it as a syncopation is not present. (It could still be argued that “stuff” is syncopated, on the grounds that its

more normative placement would be on the following quarter-note beat—since that beat is stronger—but this is debatable.) In general, stressed syllables on positions 2 and 6 tend to be less clear-cut cases of syncopation than those on positions 4 and 8.<sup>5</sup> Both misses and false positives may also result from incorrect stress values. In Figure 10C, the syllable “up” is labeled in our corpus as unstressed, since it is considered a preposition (as it sometimes is, e.g., “I walked up the stairs”); in this case, however, it is a particle, so it should be stressed. All of these problems could be addressed in various ways, but we leave this for future work; for the remainder of the study we will stick with our current measures of syncopation and anticipatory syncopation, and examine some of their applications and implications.

We have said little about weak 16th-note positions, but these deserve some discussion. Informal inspection of our corpus suggests that 16th-note syncopation is not uncommon. Figure 11A shows one example; the syllables “time” and “pain” seem to anticipate their underlying positions by one 16th-note, as shown in Figure 11B. (The syllable “-ny” might also be considered to be syncopated, but this is debatable.) It can be seen from Figure 8A that the distribution of stressed syllables on

<sup>5</sup> One reason for this is that false positives such as that shown in Figure 10B are more likely to arise on positions 2 and 6 than on positions 4 and 8. They could only arise on position 4 or 8 if there were a highly stressed syllable on 3 or 7, which seems less likely.

weak 16th-note positions (the even positions) in the RS corpus follows a similar pattern to weak 8th-note positions, with a high frequency of occurrence on positions just before a much stronger beat, such as a quarter-note beat or (even more so) a half-note beat. Again, this seems to indicate anticipatory syncopation. We can quantify the presence of 16th-note-level syncopation, as we did at the 8th-note level, by counting the number of stressed syllables on weak 16th-note beats with no syllable on or before the following strong 16th-note beat, and dividing this by the total number of stressed syllables. This yields a 16th-level SQ of .105 for the entire rock corpus, showing that syncopation at the 16th-note level is considerably less common than at the 8th-note level (which yielded an SQ of .228). We can also define an ASQ for the 16th-note level, analogous to that for the 8th-note level. In this case we count the number of syncopations (as defined earlier) on the fourth 16th of each quarter, as a proportion of the number on all weak (even-numbered) 16th-note beats:

$$\begin{aligned} 16\text{th-level ASQ} = & (S_{16}(4) + S_{16}(8) + S_{16}(12) + S_{16}(16)) / \\ & (S_{16}(2) + S_{16}(4) + S_{16}(6) + S_{16}(8) \\ & + S_{16}(10) + S_{16}(12) + S_{16}(14) \\ & + S_{16}(16)) \end{aligned}$$

The 16th-level ASQ for the RS corpus as a whole is .740. Examining ASQ values for individual songs (excluding 30 songs with denominators of zero), we find that 34 of the 50 songs have an ASQ of greater than .5, significantly more than half,  $\chi^2(1) = 41.3, p = .01$ , again suggesting a tendency toward anticipatory syncopation.

Some songs have ASQs of well below .5, both at the 8th-note level and the 16th-note level. In the case of the 8th-level ASQ, this indicates that less than half of the syncopations are at positions 4 and 8 (or at the 16th-level, less than half are at positions 4, 8, 12, and 16); this might seem to suggest the opposite tendency to anticipatory syncopation. In most such cases, however, the calculation is based on only a few notes. For example, five songs have a 16th-level ASQ of exactly zero, but in each of these songs there are only one or two 16th-level syncopations in total. (Overall, songs with nonzero 16th-level syncopations have a mean of 19.8 such syncopations.) While such songs show no evidence of anticipatory syncopation at the 16th-note level, they also show little evidence of the opposite tendency.

One could also create a measure that combined the 8th-level and 16th-level SQ's. A simple and logical way to do this would be by adding the two SQ values. This indicates the proportion of stressed syllables that are

syncopations at either the 8th or 16th levels. (There is no "double-counting" of notes here, since a note cannot be on both a weak 8th-note beat and a weak 16th-note beat.) For the RS corpus, this yields a value of  $.228 + .105 = .333$ . This is not very meaningful in itself, since there are no other corpora to compare it with (other than our small 19th-century corpus, which has an SQ of zero for both levels), but it might be useful for future comparisons. One can also combine the ASQ's for the two levels, by summing both the numerators and the denominators: the resulting value for the RS corpus is .762. However, the 8th-level and 16th-level SQs and ASQs show rather different patterns in their relationships with other variables, so we find it best to keep them separate in the following discussion.

### Further Issues and Applications

So far, our main aim has been simply to devise metrics that can be used to assess the overall levels of syncopation and anticipatory syncopation in a song or corpus. In the following section we explore some further questions regarding the use of syncopation in rock, and the possibility of answering them through corpus analysis.

A basic question to ask about the use of syncopation is, where does it tend to occur within the measure? One way to think about this is as follows. Let us assume, for the moment, that every syncopated rhythm is anticipatory—derived from an underlying unsyncopated representation. Consider a rhythm like that shown in Figure 12A: a note on the downbeat with no note on the preceding weak 8th-note beat. We will call such a note a "standard." In principle, it should be possible for any such note to be shifted to the previous weak beat (since the shift is not "blocked" by another note on that beat), producing the rhythm in Figure 12B; this is what we have defined as a syncopation. (We assume that both standards and syncopations are stressed syllables.) There are four possible locations within the measure in which a syllable on a quarter-note beat might be shifted to the previous weak 8th-note beat (see Figure 12C). We can describe each of these four locations in terms of the underlying 8th-note position of the note and the syncopated position that it may be shifted to: for example, 1→8 indicates a shift from position 1 to the previous position 8. The counts of standards and syncopations at each of the four positions are shown in Table 1. The fact that the most frequent syncopations (4 and 8) are those immediately preceding the most frequent standards (1 and 5) supports the presence of anticipatory syncopation; this is also brought out by our ASQ measure. However, there are also differences in the relative

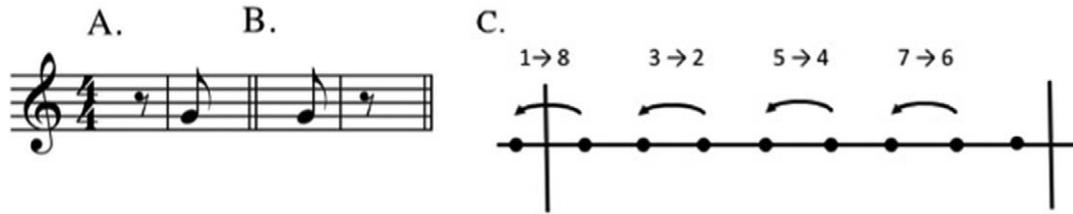


FIGURE 12. (A) Standard; (B) syncopation; (C) four possible locations for syncopation shifts.

TABLE 1. Standards and Syncopations at the 8th-note Level in the RS Corpus

Syncopation type	Standards	Syncopations	Syncs. / (Stds. + Syncs.)
1→8	1028	1046	.504
3→2	389	305	.439
5→4	639	794	.554
7→6	438	237	.351

frequency of syncopations and standards that are not explained by this view. If we examine the number of syncopations at each location as a proportion of the number of standards plus syncopations, this provides an indication of how often the syncopation option is taken, as a proportion of the number of times it *could* be taken—that is, how often a stressed syllable is shifted from the strong beat to the preceding weak one. This proportion might be described as the “syncopation tendency” of the quarter-note beat. As seen in the right-most column of Table 1, a strong pattern emerges: the syncopation tendency is strongest in the 5→4 case, then the 1→8 case, then the 3→2 case, then the 7→6 case. Chi-square tests show that the differences between all three pairs in this rank ordering are significant: 5→4 vs. 1→8,  $\chi^2(1) = 8.2$ ,  $p < .005$ ; 1→8 vs. 3→2,  $\chi^2(1) = 8.5$ ,  $p < .005$ ; 3→2 vs. 7→6,  $\chi^2(1) = 10.8$ ,  $p < .005$ . This pattern is puzzling and difficult to explain. The higher syncopation tendency values of 5→4 and 1→8 might suggest that, for some reason, syncopation shifts are more preferred at stronger beat locations; but the fact that the 5→4 syncopation is more favored than 8→1 undercuts this explanation. (As noted earlier, many of the apparent syncopations on positions 2 and 6 may not actually be syncopations, but that does not explain the current findings in any straightforward way. Indeed, if anything, this suggests that the frequency of true anticipatory syncopations on 2 and 6 is lower than suggested by the current data, which would mean that the difference between the syncopation tendencies of weak quarter-note beats and strong ones is even greater than what is shown in the table.)

Both the SQ and the ASQ show considerable variation across songs (across the possible range of 0 to 1, in both cases); how might this variation be explained? Two possible factors that come to mind are tempo and year. With regard to tempo, our 80-song corpus ranges from 58 to 234 beats per minute (BPM) (mean = 117.2, median = 114), with the majority of the songs (75%, or 60 songs) being between 60 and 140 BPM. We have observed informally that many of the songs with pervasive 8th-note syncopation are toward the faster end of this range (Chuck Berry’s “Johnny B. Goode”—167 BPM—is an example), while those with extensive 16th-note syncopation tend to be toward the slower end (the Eagles’ “Hotel California”—73 BPM—is an example). With regard to year, it is of interest to consider whether syncopation increases as the decades go by (the songs in our corpus range from 1955 to 1997). Similar questions about the effect of tempo and year can be asked about anticipatory syncopation.

Scatterplots of 8th-level SQ against tempo and year, and the same for 16th-level SQ, are shown in Figures 13A–D. (Scatterplots of ASQ against tempo and year showed little of interest, and are not included here.) There seems to be little change in 8th-level SQ as tempo increases (Figure 13A); it appears to increase slightly over time (Figure 13B). For 16th-level SQ, there is a sudden decrease above 120 BPM (Figure 13C), and there is no apparent historical trend (Figure 13D). One complication with these analyses is that year and tempo are not independent: it has been observed that the tempo of popular music has declined somewhat over the decades, from the 1960s through the 2000s (Schellenberg & von Scheve, 2012). Figure 14 verifies this phenomenon for our corpus as well; the correlation between year and tempo is moderately negative ( $r = -.54$ ). Therefore, observed changes in syncopation over the years might be partly due to changes in tempo. To address this issue, we performed a multiple logistic regression across songs, with 8th-level SQ as the dependent variable, and tempo and year as predictors; we then performed similar regressions with 16th-level SQ, and with ASQ at both 8th and 16th levels. These logistic models were

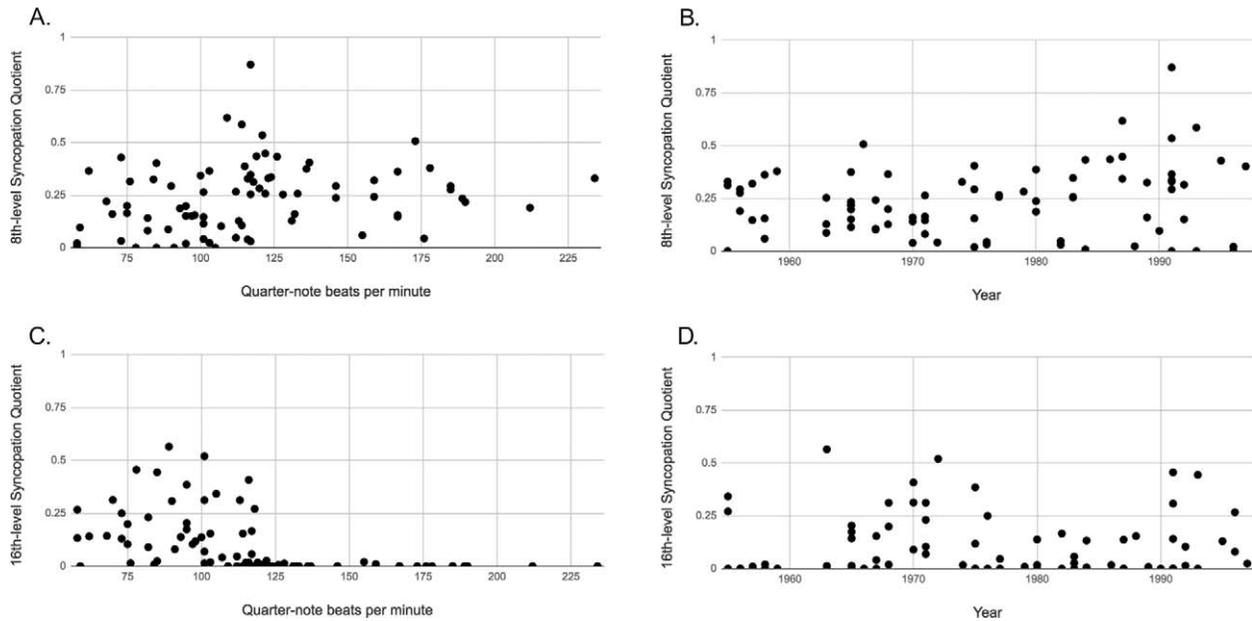


FIGURE 13. (A) 8th-level Syncopation Quotient as a function of tempo in the RS corpus. (B) 8th-level Syncopation Quotient over time in the RS corpus. (C) 16th-level Syncopation Quotient as a function of tempo in the RS corpus. (D) 16th-level Syncopation Quotient over time in the RS corpus.

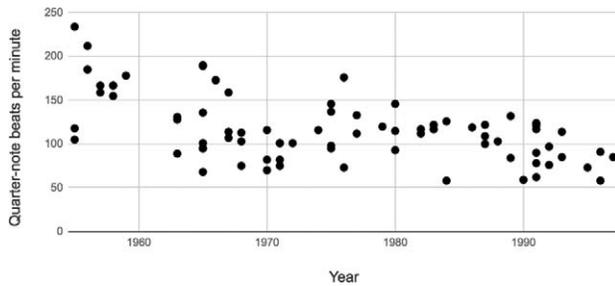


FIGURE 14. Tempo over time in the RS corpus.

then compared with models with a single predictor and with intercept-only models.

As summarized in Table 2, adding tempo and year as predictors led to statistically significant improvements in both the 8th-level and 16th-level SQ regressions compared to models in which one or both predictors were removed. In the full models, 8th-level SQ increases with tempo and year, while 16th-level SQ decreases with tempo and year. The decrease at the 16th level over time is likely driven by several slow songs from the 1990s with relatively low 16th-level SQ; while slow songs from earlier decades consistently have a high 16th-level SQ, these 1990s songs tend to feature more syncopation at the 8th level.

On the other hand, neither predictor led to a statistically significant improvement for 8th-level ASQ

compared with single-predictor or intercept-only models, and only tempo had a statistically significant effect on 16th-level ASQ. This suggests that, historically speaking, anticipatory syncopation has been a constant phenomenon in rock melodies over the years, with little apparent change at either the 8th-note or 16th-note levels. The significant effect of tempo on 16th-level ASQ may simply be a consequence of the previously observed rarity of 16th-note onsets at fast tempos: many of the faster songs in the corpus contain only a few 16th-level syncopations, and their distribution across the positions of the measure may not be very meaningful.

The effect of tempo on SQ suggests that there is an optimal duration, or range of durations, for the “unit” of syncopation (the rhythmic unit by which a syllable is shifted)—perhaps around a quarter of a second, very roughly speaking. For songs around 120 BPM, the 8th-note is closest to this optimum, but for much slower tempi, the 16th-note may be closer. We see further evidence of a distinction between the 8th and 16th metric levels in a scatterplot of 8th-level SQ against 16th-level SQ (Figure 15): there are no songs with high scores in both measures, and an overall negative correlation ( $r = -.52$ ). We should note that many songs above 120 BPM have few or no notes on weak 16th-note beats at all. London (2012) suggests that the shortest IOI for a duration to be perceived or performed as part of a rhythmic figure is 100 ms. Sixteenth notes at 120 BPM have an

TABLE 2. Results of a Multiple Logistic Regression Analysis of the Effects of Tempo and Year on 8th-level and 16th-level SQ and ASQ in the RS Corpus

A. 8<sup>th</sup>-level SQ

	Estimated coefficient	Standard error	$\Delta$ deviance with single-predictor model	<i>F</i>	<i>p</i>
Tempo (BPM)	0.011893	0.003114	-1.8391	15.068	< .001
Year	0.033670	0.009343	-1.6491	13.511	< .001

Dispersion parameter = 0.122055  
 Null deviance = 13.141 (*df* = 79), residual deviance = 10.907 (*df* = 77)

B. 16<sup>th</sup>-level SQ

	Estimated coefficient	Standard error	$\Delta$ deviance with single-predictor model	<i>F</i>	<i>p</i>
Tempo (BPM)	-0.040216	0.013253	-6.5467	47.021	< .001
Year	-0.036092	0.006077	-1.3275	9.5348	< .01

Dispersion parameter = 0.1392306  
 Null deviance = 16.4225 (*df* = 79), residual deviance = 9.8754 (*df* = 77)

C. 8<sup>th</sup>-level ASQ

	Estimated coefficient	Standard error	$\Delta$ deviance with single-predictor model	<i>F</i>	<i>p</i>
Tempo (BPM)	0.006213	0.004443	-0.56614	1.9807	.1636
Year	0.08814	0.013544	-0.12137	0.4246	.5167

Dispersion parameter = 0.2858288  
 Null deviance = 23.675 (*df* = 75), residual deviance = 23.105 (*df* = 73)

D. 16<sup>th</sup>-level ASQ

	Estimated coefficient	Standard error	$\Delta$ deviance with single-predictor model	<i>F</i>	<i>p</i>
Tempo (BPM)	-0.02771	0.00951	-3.3689	9.2428	< .01
Year	-0.02276	0.01701	-0.66175	1.8156	.1843

Dispersion parameter = 0.3644867  
 Null deviance = 23.661 (*df* = 49), residual deviance = 20.289 (*df* = 47)

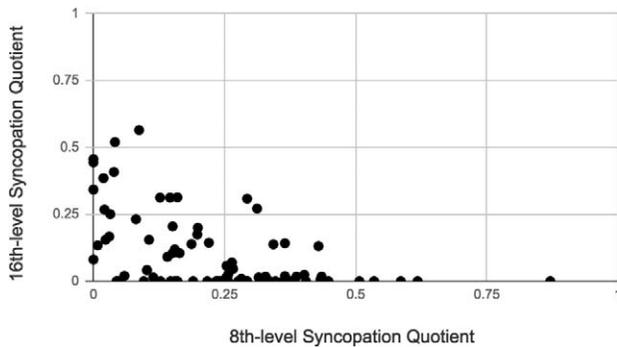


FIGURE 15. 8th-level Syncopation Quotient vs. 16th-level Syncopation Quotient in the RS corpus.

IOI of 125 ms—fairly close to London’s lower limit—which may explain the rarity of 16th-note onsets and syncopations at faster tempi. However, this reasoning

does not explain why 8th-note syncopations are *more* common at faster tempi.

It is important to remember that the identification of syncopation at the 8th-note versus the 16th-note level depends on which metrical level is chosen by the transcribers to be the main beat or “tactus” (the quarter-note level in the case of 4/4). In the original melodic transcriptions in the RS corpus, the tactus was determined by the standard rock backbeat, with kick drum on quarter-note beats 1 and 3 and snare drum on beats 2 and 4; we left this unchanged in our modified transcriptions. (In songs lacking this drum beat, or lacking drums altogether, the identification of the tactus level can be quite debatable; Bob Dylan’s “Blowin’ in the Wind” is an example in our corpus.) More recently, de Clercq (2017) has called this method of determining the tactus into question, arguing that it sometimes yields tactus levels that seem implausibly slow or

Figure 16 consists of four musical staves labeled A, B, C, and D. Each staff shows a melodic line in treble clef with lyrics underneath. Staff A is in G major (one sharp) and 4/4 time, with lyrics 'Ev - ery breath you — take'. Staff B is in G major (one sharp) and 4/4 time, with lyrics 'I — can't get no —'. Staff C is in E minor (three flats) and 4/4 time, with lyrics 'There must be some kind of way out of here'. Staff D is in E minor (three flats) and 4/4 time, with lyrics 'There must be some kind of way'. The notation in D is described as an 'ill-formed underlying representation' of the rhythm in C.

FIGURE 16. (A) The Police, “Every Breath You Take,” beginning of first verse. (B) The Rolling Stones, “(I Can’t Get No) Satisfaction,” beginning of first verse. (C) The Jimi Hendrix Experience, “All Along the Watchtower,” beginning of first verse. (D) An ill-formed underlying representation of Figure 16C.

fast—well outside the “ideal” tempo range that has been established by music cognition research (around 100–125 beats per minute). Using this criterion would double the tempo of many of the songs in the corpus, thus turning many 16th-note syncopations into 8th-note syncopations.

While a large majority of syncopations in our corpus seem explicable as anticipatory syncopations, there are a small number that—for various reasons—do not, and call out for other explanations. Three examples are shown in Figure 16. In Figure 16A, “take” is clearly stressed but is on a weaker beat than “you.” Rather than regarding this as anticipatory, it seems much more plausible to regard it as a kind of retardative syncopation or “retardation,” one that belongs on (i.e., is shifted from) the previous strong beat. (This analysis is reinforced by the fact that this unsyncopated rhythmic pattern occurs in the phrase immediately following.) One might wonder if retardations could be counted in our corpus, comparing them to unsyncopated “standards,” just as we did for anticipations. The problem is that a great many syncopations could in theory be either anticipatory or retardative—that is to say, they are both preceded and followed by empty (stronger) beats; the syllable “son” in Figure 2 is an example. Defining a retardation as a syncopation on a weak 8th-note beat with no note on the preceding 8th-note beat, we find far more of them on positions 4 (326) and 8 (526) than on positions 2 (319) and 6 (279); but there are far more retardative standards (strong-beat notes that could have been subjected to retardation but were not, i.e., those *followed* by empty weak beats) on positions 1 (1146) and 5 (814) than on positions 3 (266) and 7 (363). This would be hard to explain if the weak-beat notes were truly retardations, since the strong beats with more retardative standards (1 and 5) are followed by the weak beats with fewer

retardations (2 and 6). The explanation for this, surely, is that most of the weak-beat notes are actually anticipations, not retardations.

Another type of syncopation that does not seem convincingly explained as anticipatory is cross-rhythm (Biamonte, 2014; Traut, 2005). In Figure 16B, the four syllables of the phrase are each, essentially, a dotted-quarter note in duration, suggesting a pulse that goes against the underlying 4/4 meter. It would be *possible* to explain this rhythm as anticipatory (shifting the first syllable to the right by a quarter-note, and the second and fourth ones by an 8th-note), but we find this explanation implausible; rather, the “logic” of the melody comes from the dotted-quarter-note pulse. A final example of a non-anticipatory syncopation is shown in Figure 16C; this phrase contains a conflict between stress and meter, since “way” is more stressed than “of,” but on a weaker beat. If the phrase was heard on its own up to “way,” with a rest following, then it would be a straightforward case of anticipatory syncopation, with “way” shifted from the following downbeat; this underlying rhythm is shown in Figure 16D. However, positing this underlying representation is problematic, since there is already a syllable on the downbeat, “out,” which is clearly not syncopated (it is the main stress of the entire phrase). In this case, then, regarding the syncopation of “way” as anticipatory seems implausible. From informal inspection, syncopations like those in Figures 16A, B, and C—resisting anticipatory explanations—seem to be relatively infrequent, but they deserve further study. Lee, Brown, and Müllensiefen (2017) note that some very recent popular music contains frequent “mismatches” between stress and meter that cannot be explained away by anticipatory syncopation.

An interesting question that we have not addressed is the *function* of anticipatory syncopation, and indeed

syncopation more generally: Why do singers (and songwriters) use syncopation as they do, or at all? Three possible functions come to mind (see Temperley, 1999, for a previous discussion). First, syncopations allow the rhythm of a melody to be adjusted to fit the rhythm of the words being sung. In Figure 11, for example, the syncopated rhythm (A) seems like a much more natural setting of the words than the unsyncopated rhythm (B). Specifically, shifting the stressed syllables “time” and “pain” to the left gives them longer durations relative to the preceding unstressed syllables (“-ny” and “the” respectively). Second, shifting stressed syllables away from strong beats to preceding weak ones may prevent them from being aligned with other events in the texture—such as guitar chords and cymbal crashes (though these, too, are sometimes syncopated)—thus, perhaps, making these syllables more easily heard.<sup>6</sup> (This might explain the greater tendency toward syncopations away from strong quarter-note beats [1→8, 5→4] than weak quarter-note beats [3→2, 7→6], since instrumental events are presumably more common on stronger beats.) Third, syncopations may simply add an element of variety and complexity to the rhythmic fabric of a song—perhaps, in some cases, bringing a rhythmically simple melody up to the “optimal” level of complexity hypothesized by Berlyne (1971). The connection between syncopation and rhythmic complexity is well-established—indeed, some authors have virtually equated the two concepts, as discussed by Gómez et al. (2007). Positing syncopation as a factor in complexity seems less plausible in rock than in common-practice music, since syncopation in rock is so common; as seen in Figure 8B, some weak 8th-note positions actually have more stressed syllables than some quarter-note positions. (Complexity is often thought to be inversely related to probability, as suggested by Berlyne and others.) On the other hand, Witek, Clarke, Wallentin, Kringelbach, and Vuust (2014), focusing on percussion patterns, provide convincing experimental evidence that a moderate amount of syncopation may be optimal for the sensation of “groove.” Related to that, an issue deserving further study is the relative perceived complexity of different kinds of syncopations; it was noted earlier, for example, that anticipatory syncopations involving the shifting of unstressed syllables (like

<sup>6</sup> There may be a connection here with *melodic lead*—the tendency for melodic lines to anticipate other voices. Melodic lead has been widely observed in studies of common-practice performance, and is thought to serve a similar function, increasing the perceptual independence of the voices (Palmer, 1997). However, the timing “shifts” involved in melodic lead are very short, typically 20–50 ms (Palmer, 1997); a shift of around 250 ms is more typical of anticipatory syncopations, as discussed earlier.

Figure 10A) feel “extreme,” and therefore, perhaps, more complex.

We have argued here for the importance of syllabic stress in measuring syncopation; since syllabic stress is an important aspect of accent, incorporating it more fully captures the concept of syncopation as a conflict between meter and accent, compared to a definition based on positional factors alone. There are other sources of accent as well that could also be incorporated, such as loudness and harmonic change (Lerdahl & Jackendoff, 1983). Of course, this would make the measure more complex, and also more subjective: with all of these aspects of accent—including syllabic stress—there is some subjectivity in labeling them and also in determining their relative weight. Syncopations in rock can occur in instrumental lines (for examples, see Temperley, 1999); in that case, of course, syllabic stress is not a factor, but other non-positional sources of accent, such as dynamics, could certainly play a role.

To our knowledge, our corpus is the first publicly available corpus of melodies containing stress-annotated lyrics, and it invites exploration with regard to a number of issues beyond anticipatory syncopation. Here we consider just one. Let us use our corpus to define the “metrical strength” of each monosyllabic word. This could be done in many ways; for now, we adopt a very simple solution, which is to give a word token a strength of 1 if it occurs on a quarter-note beat or 0 otherwise; the metrical strength of a word is the average of these values across all of its occurrences. (One could also incorporate further distinctions between metrical levels, but we will not explore that here.) Under the assumption that more stressed syllables tend to be placed on stronger beats, one could regard this as a purely empirical measure of the stress level or “metrical strength” of a syllable. (For this purpose, the presence of syncopation is actually undesirable, since it means that stressed syllables often fall on weak beats; it would be preferable

TABLE 3. *The Ten Most Frequent Monosyllabic Words in the RS Corpus, with Counts and Metrical Strengths*

Word	Count	Metrical Strength
I	761	.313
the	628	.250
you	561	.335
a	399	.256
to	389	.355
and	368	.345
my	284	.352
in	283	.562
it	254	.122
me	236	.331

to use common-practice melodies, but no such corpus is available, except for the very small corpus of 19th-century English songs discussed earlier.) Table 3 shows the ten most common monosyllabic words in our corpus, along with their metrical strengths; they are all function words, and thus typically unstressed. There is considerable variation among the values, from .562 for “in” to .122 for “it”; this wide range seems to call out for explanation. One possibility is that such differences represent distinctions in the actual perceived stress level of function words—suggesting that lexical stress is a continuously varying parameter, rather than categorical, as has been assumed in the past. On the other hand, these differences in metrical strength might also be due to other factors, such as the contexts in which the words typically occur. For example, determiners generally occur before nouns, and nouns most often begin with

a stressed syllable (Kelly & Bock, 1984); this might partly account for the relatively low metrical strength of the determiners “the” and “a.” In any case, this example illustrates one kind of question that could be investigated using our corpus, and suggests that it may be a useful resource for exploring linguistic issues as well as musical ones.

### Author Note

The authors would like to thank Adam Waller for contributing code that parses the CMU Pronouncing Dictionary for computational processing, and Trevor de Clercq for helpful feedback on an earlier draft of the article.

*Correspondence concerning this article should be addressed to Ivan Tan, Eastman School of Music, 26 Gibbs St., Rochester, NY 14604. E-mail: itan@u.rochester.edu*

### References

- BERLYNE, D. E. (1971). *Aesthetics and psychobiology*. New York: Appleton-Century-Crofts.
- BERTIN-MAHIEUX, T., ELLIS, D., WHITMAN, B., & LAMERE, P. (2011). The million song dataset. In A. Klapuri & C. Leider (Eds.), *Proceedings of the 12th International Society for Music Information Retrieval Conference*. Miami, FL: ISMIR.
- BIAMONTE, N. (2014). Formal functions of metric dissonance in rock music. *Music Theory Online*, 20(2).
- BURGOYNE, J., WILD, J., & FUJINAGA, I. (2011). An expert ground truth set for audio chord recognition and music analysis. In A. Klapuri & C. Leider (Eds.), *Proceedings of the 12th International Society for Music Information Retrieval Conference*. Miami, FL: ISMIR.
- CONDIT-SCHULTZ, N. (2016). MCFlow: A digital corpus of rap transcriptions. *Empirical Musicology Review*, 11(2).
- DE CLERCQ, T. (2017). Swing, shuffle, half-time, double: Beyond traditional time signatures in the classification of meter in pop/rock music. In C. Rodriguez (Ed.), *Coming of age: Teaching and learning popular music in academia* (pp. 139–167). Ann Arbor, MI: Maize Books.
- DE CLERCQ, T., & TEMPERLEY, D. (2011). A corpus analysis of rock harmony. *Popular Music*, 30, 47–70.
- EVERETT, W. (2009). *The foundations of rock: From “Blue Suede Shoes” to “Suite: Judy Blue Eyes.”* New York: Oxford University Press.
- FOX, D. (2002). *The rhythm bible*. Van Nuys, CA: Alfred Music Publishing.
- GÓMEZ, F., THUL, E., & TOUSSAINT, G. (2007). An experimental comparison of formal measures of rhythmic syncopation. In *Proceedings of the International Computer Music Conference* (pp. 101–104). Copenhagen, Denmark.
- HALLE, J., & LERDAHL, F. (1993). A generative textsetting model. *Current Musicology*, 55, 3–23.
- HAYES, B. (1995). *Metrical stress theory: Principles and case studies*. Chicago, IL: University of Chicago Press.
- HURON, D. (2006). *Sweet anticipation: Music and the psychology of expectation*. Cambridge, MA: MIT Press.
- HURON, D., & OMMEN, A. (2006). An empirical study of syncopation in American popular music, 1890–1939. *Music Theory Spectrum*, 28(2), 211–231.
- JENNESS, D., & VELSEY, D. (2014). *Classic American popular song: The second half-century, 1950–2000*. Abingdon, UK: Routledge.
- KELLY, M. H., & BOCK, J. K. (1988). Stress in time. *Journal of Experimental Psychology: Human Perception and Performance*, 14(3), 389–403.
- KERNFELD, B. (2002). Beat. In B. Kernfeld (Ed.), *New Grove dictionary of jazz*. New York: Grove.
- KONOWITZ, B. (1991). *Alfred’s basic jazz/rock course: Lesson book, level 4*. Van Nuys, CA: Alfred Music Publishing.
- KREBS, H. (1999). *Fantasy pieces: Metrical dissonance in the music of Robert Schumann*. New York: Oxford University Press.
- LEE, C., BROWN, L., & MÜLLENSIEFEN, D. (2017). The musical impact of Multicultural London English (MLE) speech rhythm. *Music Perception*, 34, 452–481.
- LERDAHL, F., & JACKENDOFF, R. (1983). *A generative theory of tonal music*. Cambridge, MA: MIT Press.
- LONDON, J. (2012). *Hearing in time: Psychological aspects of musical meter* (2nd ed.). New York: Oxford University Press.
- LONGUET-HIGGINS, H. C., & LEE, C. S. (1984). The rhythmic interpretation of monophonic music. *Music Perception*, 1, 424–441.

- MAUCH, M., CANNAM, C., DAVIES, M., DIXON, S., HARTE, C., KOLOZALI, S., ET AL. (2009). OMRAS2 Metadata Project 2009. In K. Hirata, G. Tzanetakis, & K. Yoshii (Eds.), *Proceedings of the 10th International Society for Music Information Retrieval Conference*. Kobe, Japan: ISMIR.
- PALMER, C. (1997). Music performance. *Annual Review of Psychology*, 48, 115–38.
- PALMER, C., & KRUMHANSL, C.L. (1990). Mental representations for musical meter. *Journal of Experimental Psychology: Human Perception and Performance*, 16(4), 728–741.
- PRESSING, J. (1999). *Cognitive complexity and the structure of musical patterns*. Proceedings of the 4th Conference of the Australasian Cognitive Science Society. Newcastle, NSW, Australia: University of Newcastle.
- SCHELLENBERG, G., & VON SCHEVE, C. (2012). Emotional cues in American popular music: Five decades of the Top 40. *Psychology of Aesthetics, Creativity, and the Arts*, 6, 196–203.
- SELKIRK, E. (1996). The prosodic structure of function words. In J. L. Morgan & K. Demuth (Eds.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 187–214). New York: Psychology Press.
- SMITH, L.M., & HONING, H. (2006). Evaluating and extending computational models of rhythmic syncopation in music. In *Proceedings of the International Computer Music Conference*. New Orleans, LA: ICMC.
- STEPHENSON, K. (2002). *What to listen for in rock*. New Haven, CT: Yale University Press.
- TEMPERLEY, D. (1999). Syncopation in rock: A perceptual perspective. *Popular Music*, 18, 19–40.
- TEMPERLEY, D., & DE CLERCQ, T. (2013). Statistical analysis of harmony and melody in rock music. *Journal of New Music Research*, 42(3), 187–204.
- TEMPERLEY, N., & TEMPERLEY, D. (2013). Stress-meter alignment in French vocal music. *Journal of the Acoustical Society of America*, 134, 520–527.
- TITON, J. (1994). *Early downhome blues: A musical and cultural analysis* (2nd ed.). Chapel Hill, NC: University of North Carolina Press.
- TRAUT, D. (2005). “Simply irresistible”: Recurring accent patterns as hooks in mainstream 1980s music. *Popular Music*, 24, 57–77.
- WALLER, A. (2016). *Rhythm and flow in hip-hop music: A corpus study* (PhD dissertation). University of Rochester.
- WITEK, M. A. G., CLARKE, E. F., WALLENTIN, M., KRINGELBACH, M. L., & VUUST, P. (2014). Syncopation, body-movement and pleasure in groove music. *PLoS ONE*, 9(4) (e94446).