Corpus Studies of Harmony in Popular Music: A Response to Gauvin

TREVOR de CLERCQ [1]
Middle Tennessee State University

ABSTRACT: This paper responds to the research presented in Gauvin's paper on the evolution of harmonic syntax in popular music from the 1960s. I begin by situating the findings from his second study (on flat-side harmonies) within the context of my corpus work with David Temperley on harmony in popular music. Gauvin's results are similar to ours, although some important differences are worth noting. I also provide an interpretation of the results from his first study (on modulation), which Gauvin found to be inconsistent with his proposed hypothesis. Specifically, I conjecture that modulation and harmonic palette may be in balance with one another, in that listeners may prefer songs where harmonic content is at a medium level of complexity. Gauvin's study also brings forth some basic issues with regard to harmonic encoding, in terms of both the subjectivity of the analytical process as well as the ease by which harmonic analyses are shared with and verified by other researchers.

Submitted 2015 March 23; accepted 2015 April 9.

KEYWORDS: *popular music, rock music, harmony, modal mixture, corpus study, modulation*

IN the current issue of this journal, Hubert Léveillé Gauvin presents a corpus study that examines changes in the harmony of popular music during the 1960s, taken broadly to span 1958-1971. As Gauvin notes, musicologists generally contend that popular music became more complex during the 1960s, especially when comparing the beginning of the decade to its end. For instance, Covach (2006) argues that the Beatles evolved over the course of their career from "craftsmen" to "artists" via an increase in the complexity of the formal structures used in their songs. Gauvin investigates whether this posited increase in complexity can be found within the harmonic domain of popular music from this period as well. He does so by presenting two specific hypotheses: 1) that songs with more than one tonal center (i.e., songs that modulate) will increase in frequency over the decade, and 2) that songs featuring the flat-side harmonies of bIII, bVI, and bVII (i.e., songs that include chords other than those diatonic to a major key) will also increase in frequency. Ultimately, his results support the second hypothesis but not the first: that popular music in the late 1960s arguably may be seen as more complex, generally speaking, due to a broader harmonic palette but not through any increase in the use of modulation.

I believe the mixed results from Gauvin's study warrant some commentary. In the following paragraphs, I discuss Gauvin's article within the context of the corpus work I have conducted with David Temperley (de Clercq and Temperley, 2011). I also provide some alternative interpretations of Gauvin's findings, especially with regard to the apparent lack of support for his hypothesis on the increased frequency of modulation during this period. Finally, I consider corpus work on popular music harmony more generally, with a look toward steps that might help future research in this area.

FURTHER CONTEXT ON THE RESULTS FROM GAUVIN'S SECOND STUDY

Gauvin's article can be seen, at least to some extent, as a follow-up to the corpus study of rock harmony that Temperley and I published in 2011 (de Clercq and Temperley, 2011). As Gauvin mentions, we reported in our 2011 article that flat-side harmonies (e.g., bIII, bVI, and bVII) were rare in songs from the 1950s and were more common in songs from the 1960s and onwards (see Table 1). It is thus not wholly surprising, given our findings, that Gauvin's results show a similar increase in the use of flat-side harmonies from the early 1960s to the late 1960s. That being said, Gauvin's research extends and complements ours in a few different ways that are worth highlighting.

Table 1: Overall proportion of chromatic roots in each decade, 1950s-1970s (an excerpt from Table 8, originally published in de Clercq and Temperley, 2011).

	1950s	**1960s**	**1970s**
I	0.423	0.327	0.302
bII	0.001	0.000	0.004
II	0.004	0.074	0.038
bIII	0.000	0.009	0.032
III	0.007	0.040	0.027
IV	0.321	0.239	0.226
#IV	0.000	0.001	0.009
V	0.221	0.146	0.154
bVI	0.001	0.003	0.054
VI	0.011	0.072	0.063
bVII	0.007	0.084	0.089
VII	0.006	0.006	0.004

One important aspect of Gauvin's work is that he investigates harmonic changes over a narrower time span, i.e. the 1960s, whereas my study with Temperley spanned the 1950s through to the 1990s. The narrower historical window in Gauvin's study allows for a more refined speculation on the historical ascendancy of these specific chords. Temperley and I conjectured that the increased prominence of flat-side harmonies reflected the rise of hard rock and heavy metal (p. 64), since bIII and bVI are used only very rarely until the 1970s. Yet very few songs in Gauvin's corpus can be considered to be hard rock or heavy metal. Given Gauvin's results, it may be that the increased use of these harmonies began earlier, sometime in the mid- to late-1960s, and initially arose from more pop-oriented styles. Speculation aside, Gauvin does provide evidence that confirms the basic premise that harmony in popular music from the 1950s and a good portion of the 1960s does generally avoid flat-side harmonies. It is also helpful to see our findings supported with a much larger selection of songs, since the results from de Clercq and Temperley (2011) derived from a corpus with only 20 songs per decade.

It is also worthwhile to see our findings reproduced in an different body of songs. The corpus that Temperley and I compiled was comprised of songs that achieved critical acclaim, drawn from *Rolling Stone* magazine's 2004 list of the "500 Greatest Songs of All Time." In contrast, Gauvin's corpus is comprised of songs that achieved commercial success, culled from the *Billboard* Hot 100 charts. It is hard to say which rubric – critical acclaim or commercial success – gives a better picture of a decade. While the *Billboard* charts are one useful snapshot, many widely popular songs that received extensive radio airplay during their day such as "Stairway to Heaven" by Led Zeppelin never reached the *Billboard* charts because they were never officially released as a single. Any corpus attempting to represent popular music in general is inherently limited. Accordingly, evidence for a phenomenon that is found within multiple corpora is of great value.

One other difference worth noting is that Temperley and I reported chord root proportions overall while Gauvin reports chord incidences on a song-by-song basis. For example, Gauvin finds that 28.2% of songs in the late 1960s include at least one instance of a bIII chord, as compared to only 10.7% of songs in the early 1960s (p. 8). In contrast, our results (reproduced in Table 1) showed that only about 0.9% of chords overall in songs from the 1960s are built on scale-degree b3. Let us assume, for the sake of argument, that these statistics are roughly representative of the larger population. It would imply that, while songs began to use the bIII chord much more frequently in the late 1960s, the bIII chord still played a relatively small role in the harmony of each individual song. I cannot say that this is necessarily true, of course, since I am only speculating from the given results but I can say that Gauvin's data complements my work with Temperley and offers some potential avenues for future research along these lines.

AN INTERPRETATION OF THE RESULTS FROM GAUVIN'S FIRST STUDY

Although the results from Gauvin's study of flat-side harmonies support his general hypothesis of increased harmonic complexity during the 1960s, the results from his study on modulation do not. The issue of whether or not harmonic complexity increases during this era thus appears to be unresolved. Gauvin does not offer an explanation in his article, so I would like to propose one possible explanation here.

To put it simply, I posit that modulation may exist in balance with an increased harmonic palette (at least with regard to popular music from this era). In other words, as a greater variety of

chords are introduced into the harmonic palette, modulation becomes less likely. Conversely, as a smaller variety of chords are used in a song, modulation becomes more likely.

My conjecture derives from the two-factor arousal theory of Berlyne (1971), wherein certain properties, such as complexity, increase arousal while other properties, such as familiarity, reduce arousal. Berlyne theorizes that aesthetic preference will peak with moderate levels of arousal. So, for example, too much or too little complexity will result in a relative decrease in aesthetic preference. The relationship between complexity and preference can be represented through an inverted-U curve or "Wundt curve", as shown in Figure 1. Experimental support for the inverted-U relationship between subjective complexity and preference can be found in a number of experimental studies (e.g., Hargreaves, 1984; Orr and Ohlsson, 2005; Tan, Spackman, & Peaslee, 2006).

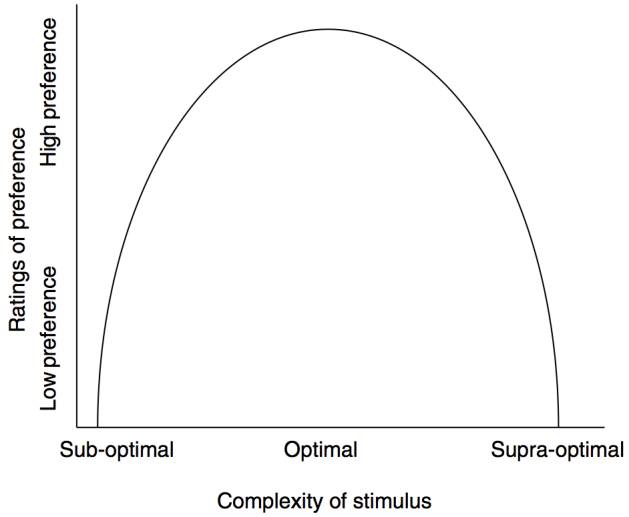

Figure 1: The inverted-U relationship between complexity and preference, as theorized in Berlyne (1971). Reproduced from Tan, Spackman & Peaslee (2006).

The trade-off between complexity and preference within the harmonic domain can also be framed in terms of "communicative pressure" (Temperley, 2004) or a "trading relationship" (Swain, 1997). It could be said that the introduction of a greater variety of chords in the harmony of a song causes the tonal center to destabilize, at least to some extent. Since modulation also serves to destabilize the tonal center, the combination of both chromatic or modal harmony and modulation might result in confusion about (or at least a serious weakening of) the tonal center. Perceived complexity is a subjective domain and it is possible that expert listeners have a higher optimal level of complexity. However, popular music, unlike art music, is inherently constrained by the listening preferences of the general public. So, while high levels of chromatic/modal harmony and modulation may generate high preference ratings in a certain segment of listeners, the average listener may prefer a more moderate level of complexity within the harmonic domain.

It is worth pointing out that Gauvin's data can, in fact, be seen to verge on support for this optimal harmonic complexity theory. In his Figure 2.1, there appears to be a general decrease in multi-tonic songs from the early 1960s to the late 1960s, which mirrors the increase in flat-side harmonies during this period. This decrease turns out to not be statistically significant, although the *p*-value of 0.518 does border on significance. With some re-analysis of Gauvin's data, it is possible to present statistics that would have shown a significant decrease over time in songs that modulate. For example, consider Table 2 below, which presents the raw data from Gauvin's study. (This table was generated using the information provided in Gauvin's Figure 2.2 and Table 1.)

Table 2: Distribution of multi-tonic songs per year.

Year	Number of Songs Total	Number of Multi-Tonic Songs	Multi-Tonic Songs (as proportion of total)
1958	7	1	0.143
1959	14	2	0.143
1960	9	1	0.111
1961	20	1	0.050
1962	27	5	0.185
1963	20	3	0.150
1964	25	5	0.200
1965	20	1	0.050
1966	18	0	0.000
1967	28	3	0.107
1968	26	1	0.038
1969	24	3	0.125
1970	22	2	0.091
1971	32	3	0.094

If we discount the beginning and end years (1958 and 1971), we are left with a twelve-year span (from 1959-1970) that can be divided up into three 4-year timespans: the early 1960s (1959-1962), the mid 1960s (1963-1966), and the late 1960s (1967-1970). In Table 3, the number of multi-tonic songs in the corpus from the early, middle, and late parts of the broad decade remains constant while the number of songs total in the corpus for each timespan increases. As a result, the percent of songs per four-year period shows a consistent decrease. If we take the average year within each timespan as the value for the timespan as a whole, the correlation between year and percent of multi-tonic songs turns out to be highly significant ($r = -1.00$, $df = 1$, $p < .0001$).

Table 3: Distribution of multi-tonic songs per four-year period.

Year	Number of Songs Total	Number of Multi-Tonic Songs	Multi-Tonic Songs (as percent of total)
1959-1962	70	9	0.129
1963-1966	83	9	0.108
1967-1970	100	9	0.090

I provide this information not to prove anything, per se; this sort of *post-hoc* manipulation goes against the basic premise of a *p*-value, i.e., making a hypothesis and then testing it. Huron (2013) warns specifically against relying too heavily on these sorts of inferences. Nonetheless, this result does suggest a research question that could be tested in future work: Do popular songs that include a smaller variety of chords more often involve modulation, and vice versa? Of course, modulations, such as those between relative keys, can often be ambiguous, and this is especially true for popular music in my experience. So there are some preliminary issues that would have to be addressed. Moreover, one would have to use a different corpus than the *Billboard* set to test this question, since it would be poor experimental design to re-test the same corpus that engendered this new hypothesis. This leads me to my last point.

ON ENCODING AND ENCODING FORMATS

Corpus work on harmony in popular music is important, if only because it allows us to test some of the claims made in the music theory literature about the nature of harmony in this style. Stephenson (2002), for example, claims that the standard root movement of rock harmony operates by descending seconds, ascending thirds, and descending fourths, in contrast to the standard root movement of common-practice harmony, which operates by ascending seconds, descending thirds, and ascending fourths. Stephenson's claim could be easily assessed via a corpus study that would analyze the distribution of root movements in rock music versus those in common-practice music. It is doubtful that there will ever be enough songs (or enough corpora of songs) to test and prove all of the claims made in the music theory literature, much less all of the interesting questions that we as researchers hope to answer about this music. The primary hurdle is that the encoding process is fairly laborious and time-consuming, and so the data with which we have to work is relatively limited. It was, at least in part, for this reason that Burgoyne and his colleagues at McGill created the *Billboard* data set

(Burgoyne, 2011; Burgoyne, Wild & Fujinaga, 2011): to act as a "ground truth" that could train computer algorithms to correctly identify chords within streams of raw audio. In short, if a computer can accurately turn an audio file into a sequence of chords, then we can simply let it loose on our iTunes library and generate statistics from the output. This is a noble goal, and I expect audio chord recognition algorithms will continue to improve in accuracy and quality in the coming years. There will always be, however, an element of harmonic analysis that is subjective or ambiguous. As Temperley and I reported (de Clercq & Temperley, 2011), we found that we agreed on the chromatic relative root of a chord (e.g., is it a tonic or dominant?) only about 92% of the time.

To highlight these and other problematic aspects of harmonic analysis, I would like to offer some alternative chord choices for a few different parts of the excerpt of "Sidewalk Surfin'" by Jan & Dean that Gauvin uses as his example of the *Billboard* encoding format (shown in his Figures 1.1 and 1.2). The first alternative involves the opening bars, which have been encoded as having no harmony. Indeed, the song arrangement here has no instrumental support, only vocals. But the vocal harmonies can also be viewed as implying the | V bIII | IV V | turnaround that occurs later in the song (at 0:35), and so it is reasonable to say that these bars do indeed contain harmony. Another ambiguous analytical situation occurs later in the song, one bar before the modulation to Eb major (around 1:08). Here, the Eb7 chord has been encoded as a bII7 chord by Gauvin's algorithm (since the original transcriptionists encoded the modulation as occurring only after this chord), but I could also hear this Eb7 chord as the tonic of the new key (embellished with a bluesy dominant seventh). In other parts of the song, I hear different absolute chord roots than those notated in the transcription. The last bar of the first verse (around 0:40) sounds more to me like a continuation of the D major chord with a major seventh on top than an A major chord. The first two bars after the modulation to Eb major (around 1:10) sound more to me like two bars of a Bb major chord (V) than two bars of an Eb major chord (I).

I do not want to make too much of these analytical differences, especially since Gauvin is simply using a pre-existing set of transcriptions. Moreover, most of the chords in this "Sidewalk Surfin'" excerpt reflect my own hearing of the song and are perhaps entirely unambiguous. Presumably, the relatively small percentage of ambiguous or subjective (or incorrect) harmonic content in each song becomes statistical background noise once it becomes part of the larger corpus. It is for this reason that corpus studies of harmony in popular music would benefit from having multiple researchers encode the same songs to help dilute and thereby minimize human error as well as simply having more songs encoded.

To that end, work in this field would be aided greatly by a standardized encoding format, so that researchers could share transcriptions as well as the computer programs used to parse those transcriptions. The *Rolling Stone* corpus that Temperley and I created uses one encoding format for harmony and form: a recursive system with traditional Roman numerals; the *Billboard* corpus uses a different one: a non-recursive system with standard pop notation; the Centre for Digital Music at Queen Mary University of London uses yet another format (see Harte, 2010). Other encoding formats can easily be imagined, if they are not already in use. Each format has its advantages and disadvantages, which are not worth discussing here. It is not impossible, of course, to convert and translate between different encoding formats; however, it makes this type of research much more difficult. For example, some considerable portion of Gauvin's work for his article was spent converting the existing *Billboard* transcriptions from a pop chord format to a function-based chord format. I will not try to propose a standardized harmonic encoding format here; such a proposal would be beyond the scope of my commentary. But now that this type of research is becoming more common, it is time for us as a music research community to think seriously about how this type of work can be shared and corroborated more easily.

NOTES

[1] Correspondence concerning this article should be addressed to: Trevor de Clercq, Department of Recording Industry, 1301 East Main Street, Box 21, Middle Tennessee State University, Murfreesboro, TN 37132, USA, trevor.declercq@mtsu.edu.

REFERENCES

Berlyne, D. E. (1971). *Aesthetics and psychobiology*. New York: Appleton-Century-Crofts.

Burgoyne, J. A. (2011). *Stochastic processes and database-driven musicology*. (Unpublished doctoral dissertation). McGill University, Montreal.

Burgoyne, J.A., 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* (pp. 633-638). Miami, Florida: University of Miami.

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

Everett, W. (2004). Making sense of rock's tonal systems. *Music Theory Online,* 10(4). Retrieved from http://www.mtosmt.org/issues/mto.04.10.4/mto.04.10.4.w_everett.html

Gauvin, H. L. (2015). "The times they were a-changin'": A database-driven approach to the evolution of harmonic syntax in popular music from the 1960s. *Empirical Musicology Review*, 10(3), 215-238.

Hargreaves, D. J. (1984). The effects of repetition on liking for music. *Journal of Research in Music Education*, 32(1), 35-47.

Harte, C. (2010). *Towards automatic extraction of harmonic information from music signals*. (Unpublished doctoral dissertation). University of London, London.

Huron, D. (2013). On the virtuous and the vexatious in an age of big data. *Music Perception*, 31(1), 4-9.

Moore, A. (2001). *Rock: The primary text*. Aldershot, UK: Ashgate Press.

Orr, M. G., & Ohlsson, S. (2005). Relationship between complexity and liking as a function of expertise. *Music Perception*, 22(4), 583-611.

Stephenson, K. (2002). *What to listen for in rock: A stylistic analysis*. New Haven: Yale University Press.

Swain, J. (1997). *Musical languages*. New York, NY: W. W. Norton & Company

Tan, S., Spackman, M. P., & Peaslee, C. L. (2006). Effects of repeated exposure on liking and judgments of thematic unity of patchwork and intact compositions. *Music Perception*, 23(5), 407-421.

Temperley, D. (2004). Communicative pressure and the evolution of musical styles. *Music Perception*, 21(3), 313-337.