# Rhythmic Complexity in Jazz:
# An Information Theory Approach

*Revised Version*

Douglas R. Abrams

Based on the doctoral dissertation
successfully submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
February 3, 2023

## ACKNOWLEDGEMENTS

I would like to thank my dissertation advisor, Chris White, for his constant availability and efforts to assist me in the completion of this research.

I would also like to thank Michael Lavine, Professor Emeritus of Math and Statistics at UMass Amherst, and Anna Liu, Professor of Math and Statistics at UMass Amherst, for their assistance with the statistics used in this project. I would like to thank Luke Remage-Healey, Professor of Psychological and Brain Sciences at UMass Amherst, for enlightening conversations regarding statistics. And I would especially like to thank Yujian Wu, Ph. D. student in the department of Math and Statistics at UMass Amherst, for his constant availability and readiness to help with the statistics used in this project. All statistical errors are, of course, my own.

# ABSTRACT

Information theory can be applied wherever measurements are describable by random variables. Here I applied it to the study of jazz rhythm, using entropy to study perceived rhythmic complexity, to ascertain rhythmic differences between soloists, and to quantify the increase in complexity in embellished versions of popular song, and using mutual information to study soloist-accompanist interactions.

I experimentally studied the effects of entropy, periodicity, syncopation, number/density of notes, and order effects on perceived rhythmic complexity. One model showed that entropy, number/density of notes, and order effects were significant factors in the perception of rhythmic complexity; another showed that periodicity, syncopation, number/density of notes, and order effects were significant. Thus entropy is a significant factor in the perception of rhythmic complexity, though it is likely that it is mediated by periodicity.

I analyzed eighty-eight transcriptions of solos by Armstrong, Hawkins, Young, Christian, and Parker, and made pairwise comparisons for entropy. I found that Young's solos had significantly greater entropy than solos by Armstrong, Christian, and Parker. One pairwise comparison, Hawkins vs. Parker, was surprising in that the entropy for Parker was less (though not significantly so) than that for Hawkins. One possible explanation is that the density of stress accents for Hawkins was greater than that for Parker, while the opposite was true for contour accents. I also used the estimated marginal means technique to debunk two commonly held chronological notions about the solos of Armstrong and Young.

Finally, I used mutual information to study soloist-accompanist interactions. I did so by calculating MI for accented notes in ten Parker solos with Charleston comping rhythms, actual comping rhythms and random comping rhythms. MI was highest for random rhythms and lowest for Charleston rhythms.

## 1 Introduction

The field of research on rhythmic complexity is a fertile one. In recent decades, a plethora of methods for quantifying rhythmic complexity has been introduced. Nonetheless, an exact definition of rhythmic complexity is hard to come by. One intuitive definition might be that rhythmic complexity reflects the difficulty of performing a given rhythm. Another might reflect the difficulty of tapping along with a rhythm, or recalling a rhythm after a certain amount of time has elapsed. And in some cases, a listener might *imagine* how difficult it would be to perform a particular rhythm. In fact, each of these definitions has been used, e.g. in studies by Shmulevich and Povel (2000), and Fitch and Rosenfeld (2007). In any case, the approach taken here asked experimental subjects to rely on their own intuitive ideas about rhythmic complexity in order to rate excerpts for complexity on a scale of one to seven.

The study of rhythmic complexity allows us to compare complexity in music to complexity in other fields. This is partially due to the fact that at its most basic, music can be understood as a form of communication. Semiotically speaking, of course, it is more vague than other forms of communication such as speech – the relationships between signifier and signified are less specific in music than in speech – but it is useful nonetheless to treat music as communication, for reasons discussed below. Though in most cases it is unlikely that any two different listeners will agree 100% on *what* a given piece of music communicates, they might agree more often on *how* it communicates – and the analysis of *how* a piece communicates is the realm of music theory.

Beyond speech and music, of course, there are many other forms of communication as well, notably the digital transmission of sound or video; the theory undergirding such types of communication dates to the mid-20th century, and goes under the name of *information theory*. Shannon (1948) laid the groundwork for this field in his seminal work *A Mathematical Theory of Communication*. Since then, information theory has been applied in many fields outside of communication theory, including music.

How is it possible that a theory devised for the electrical transmission of signals can also be applied to music? The answer comes from Leonard Meyer and his work linking expectation in music – specifically, *thwarted* expectation – to meaning (1956, 1957). Meyer posits that low probability events in the context of a given piece (that is, events which thwart our expectations) have more information, and thus more meaning, than high probability events. Keeping in mind that, quoting Joel Cohen (1962), "statistical probability, or relative frequency, corresponds to the listener's expectations," there is then a natural correspondence between information theory and meaning in music. To wit, information theory deals with Cohen's probabilities – thus, with expectation – and these comprise the *information content* of a set of events; hence the correspondence between music theory and information theory.

Thus, treating music as a form of communication allows us to apply the techniques of information theory to music. That is what is done here.

### 1.1 Introduction to Entropy

Shannon Entropy (here, just "entropy"), introduced in *A Mathematical Theory of Communication*, is information theory's most fundamental concept, and has been employed in many fields. Roughly speaking, it measures the "unpredictability" of a sequence of events (Cover and Thomas, 2006). Its wide applicability is probably due to its simplicity: as will be demonstrated later, any quantity that can be described by a probability distribution can be described, using a simple formula, by entropy. As a quick Wikipedia search will indicate, fields in which entropy has been applied include statistical inference, cryptography, neurobiology, perception, bioinformatics, thermal physics, quantum computing, information retrieval, intelligence gathering, plagiarism detection, pattern recognition, anomaly detection, and others. Therefore, it should not come as a surprise that it has been applied in the field of music theory.

A cluster of papers applying entropy to music appeared in the fifties and early sixties (Pinkerton 1956; Meyer 1956, 1957; Youngblood 1958; Cohen 1962), applying the concept of entropy to pitch structures and laying the philosophical groundwork for the field. A PhD dissertation in 1959 (Crawley), which I have not examined, pioneered the application of information theory to rhythm, and a single paper applying information theory to both pitch and rhythm (Hiller and Fuller) appeared in 1967. Several papers in the eighties and early nineties refined the earlier work (Knopoff and Hutchinson 1981; Knopoff and Hutchinson 1983; Snyder 1990), and since 2002, there have been a number of papers and books on information theory and music, going beyond entropy-based analysis and working with the information content of individual events and with a generalization of entropy called Markov chains: Abdallah 2002; Pearce and Wiggins 2004; Sadakata et al. 2006;.Huron 2006; Temperley 2007; and Margulis and Beatty 2008. Finally, since the late nineties and early two-thousands, a number of papers using Markov chains to analyze and compose *jazz* have appeared, including: Yarom (1997); Gillick, Tang, and Keller (2010); Pfeiderer and Krizler (2010); Bell (2011); Victor (2012); Norgaard, Spencer, and Montiel (2013); Linskens and Schoenmakers (2014); Yun (2016); Rouse (2017); Quick and Thomas (2019); and Frieler (2019).

While experiments have been used to compare many computed measures of rhythmic complexity to perceived rhythmic complexity, to my knowledge, experiments directly involving entropy have been rare: an article by Thul and Toussaint (2008), and two by De Fleurian et al. (2014 and 2017). And to my knowledge, there have been no previous studies of mutual information applied to music.

## 1.2 Mathematical Preliminaries I: Definition of Entropy

A "random variable" is a variable which, when measured, can take on one of several values with a given probability of occurrence for each value[1]. The set of values a random variable can take on is called its "sample space," and the set of probabilities with which a random variable takes on its various values is described by its "probability distribution." The probability with which a given value of the random variable occurs ranges from zero (no probability of occurring; this value would not technically be a part of the probability distribution) to one (100% probability of occurring; this value would be the only value in the probability distribution). The sum of all values represented by a probability distribution is always one, since *something* must occur.

Most of the random variables we will be working with measure the inter-onset intervals (IOIs) between dynamically accented notes in solos by well-known jazz

[1] I will focus here on "discrete" random variables, since the calculated entropy values I use are constructed from discrete random variables. This is not affected by the fact that the entropy values themselves are treated as continuous variables.

musicians.[2] Roughly speaking, dynamically accented notes are simply notes which are louder than surrounding notes (Roeder 2017); a more detailed explanation of how accented notes are identified is given later, in the section on computational results. Inter-onset intervals are defined as the durational distances between the attacks or onsets of pairs of notes (Patel and Danielle 2003). While the time delays between accented notes in a solo are not, of course, actually *random*, we can glean important information about a solo (in particular, about the unpredictability of its rhythms) by treating it as though time delays have been selected at random from the distribution of time delays describing that solo.

Figure 1 provides an example of how I obtain probability distributions from solos; there are other ways, of course, but this is the method I use. Figure 1 is a brief excerpt from Lester Young's solo on "Blues for Greasy;"[3] accent marks indicate dynamic accents. Having identified dynamically accented notes, the next step is to tabulate the inter-onset intervals.

**Figure 1.** Musical example for calculating a probability distribution

Note that for solos in which the eighth-notes are "swung," care must be taken to identify when an accent falls on the long or short part of the beat, since this may affect the IOIs involving that accent. In this example, there are two examples of IOIs that

[2] In Section 5, the random variables are the sixteen possible combinations of eighth-note onsets and ties or rests in a two quarter note window of time.
[3]Note that this was almost certainly not recorded as a "blues in B;" it was most likely a blues in C or B flat, and the discrepancy is probably due to problems with the recording apparatus. This does not affect entropy.

depend on whether they start with long eighths or short eighths: seven eighth-notes starting on the long part of the beat (beat one in m. 2 to the “and” of four in m. 2; beat one in m. 5 to the “and” of four in m. 5), and seven eighth-notes starting on the short part of the beat (the “and” of one in m. 4 to beat one in m. 5). A distinction must be made since one of them will have duration L+S+L+S+L+S+L and the other will have duration S+L+S+L+S+L+S, where L is the duration of the long half and S is the duration of the short half of the swung eighth-note pair. (The distinction need not be made for IOIs consisting of an even number of eighth-notes).

We can assign numerical values to each of these IOIs, assuming that the ratio of the long part of the beat to the short part of the beat is 2:1, in other words, that the length of the long part of the beat is 1.333 eighth-notes and the length of the short part of the beat is 0.667. (This assumption will be examined more closely in Section 3). Note that the ratio is 2:1 and that they sum to two eighth-notes, or one quarter. The IOI of seven eighth-notes starting on the long part of the beat is 7.333 eighth-notes long, while the IOI of seven eighth-notes starting on the short part of the beat is 6.667 eighth-notes long.

With this in mind, the frequency of each IOI may be tabulated as follows: two eighth-notes, one; 7.333 eighth-notes, two; 6.667 eighth-notes, one; ten eighth-notes, one. Note that there are six accents, thus five IOIs. Probabilities are obtained by dividing frequencies by five. The resulting distribution is given in Table 1.

| **IOI (eighth-notes)** | 2.00 | 6.67 | 7.33 | 10.00 |
|---|---|---|---|---|
| **Probability** | 1/5 | 1/5 | 2/5 | 1/5 |

**Table 1.** Probability distribution for excerpt in Figure 1

In his 1948 paper, *A Mathematical Theory of Communication*, Claude Shannon derived the following formula for entropy based on three simple axioms:

$$H \equiv -\sum_{x} p(x)\log(p(x)) \qquad \text{Eq. 1}$$

where *p(x)* is the probability with which a variable *X* takes on the value *x* and the sum is over all possible values of x. The convention adopted here, common in information theory, will be to use logarithms of base 2, so that entropy will be measured in what information theorists call "bits."

Continuing with the current example, we can calculate entropy directly from Equation 1 using the probability distribution in Table 3.

The calculation is given by:

$$H = -(1/5)\log_2(1/5) - (1/5)\log_2(1/5) - (2/5)\log_2(2/5) - (1/5)\log_2(1/5) = \qquad \text{Eq. 2}$$

$$(3/5)*\log_2(5) + (2/5)\log_2(5/2) = 1.922$$

While this is an artificially simple example, it captures the essence of how I calculate entropy for musical excerpts. Note, by the way, that the entropy calculation does not explicitly rely on the values of the IOIs; all that matters is the form of the probability distribution itself.

### 1.3 Mathematical Preliminaries II: Further Examples and Caveats

Here I examine three contrasting examples.

First, consider the rhythm shown in Figure 2, which represents an excerpt in which the first of every four beats is accented (only accented notes are displayed). This is not an uncommon situation in unsyncopated music. Intuitively, we expect the entropy to be zero, since there is no uncertainty in the IOI's. Mathematically, there is only one non-zero probability, and it must equal one. Therefore, the logarithm in equation 2.1, and thus the entropy, is equal to zero.

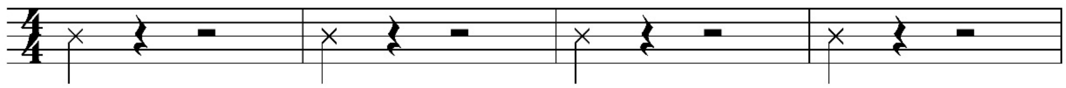

**Figure 2.** A zero entropy rhythm

Next consider the situation illustrated in Figure 3; there is considerably more variation in time delays here. First of all, there are eight different time delays represented here, from one eighth note to an entire measure. Secondly, each of these values occurs with equal probability (1/8). This situation (equal probability for each value) represents the maximum entropy distribution for a given number of possible values a random variable can take on. For *n* possible values, the maximum entropy is $\log_2 n$, and in this case the entropy is $\log_2 8 = 3$.

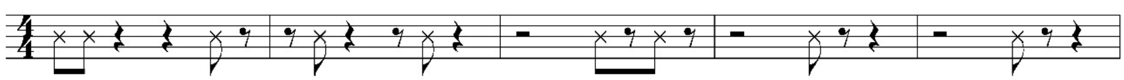

**Figure 3.** A maximum entropy rhythm

Finally, consider the situation illustrated in Figure 4, the Charlie Parker composition "Billie's Bounce"; the distribution of probability of occurrence vs. IOIs is shown in Figure 5. As discussed previously, care must be taken to identify these IOIs; though at first glance it might appear that there is only one IOI – three eighth-notes – between the first seven pairs of accents, some of them start on long eighth-notes and some of them on short eighth-notes. So in Figure 5, note that there are points that fall just to the left of three on the x axis and points that fall just to the right of three on the x axis, reflecting the unevenness of the swung eighth-note pairs. The IOI between the accent on beat four of measure three and the accent on beat one of measure four is two eighth-notes, between the accent on beat one of measure four and the accent on beat three of measure four is four eighth-notes, and so on.

The entropy of this melody is 2.503, and it has seven distinct IOIs. The entropy for seven equiprobable categories of IOI (the maximum-entropy situation for seven IOIs), would be $\log_2 7 = 2.807$. It is not surprising that the entropy for “Billie’s Bounce” is less than the entropy for eight equiprobable IOIs (the example discussed above), since, a) there are fewer IOI categories for “Billie’s Bounce,” and b) the distribution is not equiprobable (see Figure 5), a requirement for maximum entropy given a fixed number of distinct IOIs.


**Figure 4.** “Billie’s Bounce” (head) by Charlie Parker

Some authors define a quantity I will call “normalized entropy,” which I define here in the context of IOIs and rhythm. Normalized entropy, for a rhythm having *n* distinct IOIs, is obtained by dividing entropy by the maximum entropy obtainable for that many IOI’s: $\log_2 n$. Thus, for *n* distinct IOI’s, and unnormalized entropy *H*, normalized entropy is given by the following expression:

$$H_N = H / \log_2 n \qquad \text{Eq. 3}$$

Some authors use the term "relative entropy" for this quantity, but this is a misnomer since, according to Cover and Thomas (2006), the term "relative entropy" is reserved for something else (see below).

While normalized entropy gives an idea of how high the entropy of an excerpt is relative to its maximum value, it has a serious flaw. To see this, consider the probability distribution for the Charleston rhythm (rhythm shown in Figure 6). It consists of two IOIs: 3 eighth notes and 5 eighth notes, both of which occur once in each measure. Thus, the entropy is maximized for this many distinct IOIs, and the normalized entropy is:

$$H_N = (-0.5*\log_2(0.5) - 0.5*\log_2(0.5))/\log_2(2) = \qquad \text{Eq. 4}$$

$$0.5*\log_2(2)+0.5*\log_2(2) = 1$$

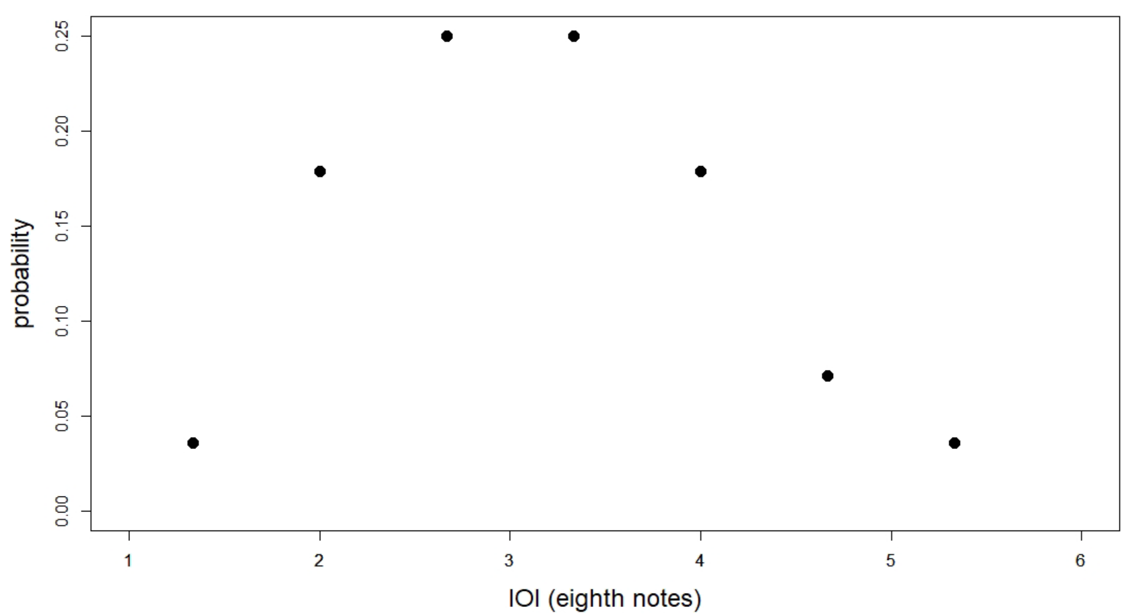


**Figure 5.** Probability distribution for Parker's "Billie's Bounce"

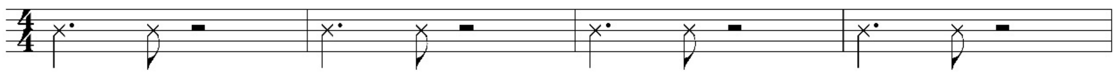

**Figure 6.** The Charleston rhythm

A much more complex rhythm, however, might have substantially lower normalized entropy. For example, the melody of "Billie's Bounce" would have a normalized entropy of just 0.892, since the raw entropy obtained from it is 2.503, and the entropy for seven equiprobable categories is 2.807. Margulis and Beatty (2008) call this the "rare interval problem": if one adds just one instance of an IOI not already present in a given distribution, the denominator of the normalized entropy formula increases substantially, while the unnormalized entropy increases only slightly and this reduces the normalized entropy.

Despite the problems inherent in using normalized entropy, many authors use it to define a related concept, redundancy, using the following formula:

$$R = 1 - H_N \qquad \text{Eq. 5}$$

Redundancy varies between zero and one, with zero corresponding to maximum normalized entropy, and one to minimum normalized entropy. Intuitively, $R$ represents the predictability of a probability distribution, though given the limitations of normalized entropy, it must be used with caution.

In Chapter 5 it will be shown that the issue of normalization can be sidestepped by evaluating the entropy for each excerpt separately, and making pairwise comparisons between musicians based on entropy using a method called "estimated marginal means."

## 1.4 Mathematical Preliminaries III: Relative Entropy and Mutual Information

Having discussed entropy, normalization, and redundancy, I now introduce the concepts of relative entropy and mutual information.

The "relative entropy," also called "Kullback-Leibler distance" or "Kullback-Leibler divergence" (I will use the abbreviation KLD) of two random variables defined on the same sample space with probability distributions $p(x)$ and $q(x)$ is given by (Cover and Thomas, 2006):

$$D(p \| q) \equiv \sum_{x} p\,(x) \log_2(p(x)/q(x)) \qquad \text{Eq. 6}$$

This is not a true "distance," since, for one thing, it is not symmetrical with respect to $p$ and $q$, but it is often useful to think of it as one (ibid.). It will be used here to define mutual information.

Several terms must be defined in order to understand mutual information. The elements of the product space $x \otimes y$ are all of the ordered pairs $(a, b)$, where $a$ is drawn from the set comprised of all values of $x$ and $b$ is drawn from the set comprised of all values of $y$. Given a random variable $X$ and a random variable $Y$, their joint probability distribution, defined on the product space $x \otimes y$ and notated $J(x,y)$, is the probability that $X = x$ and $Y = y$ simultaneously. The marginal distributions, $p(x)$ and $q(y)$, give the probabilities that $X = x$ and $Y = y$ individually. The marginal distributions can be calculated from the joint distribution as follows:

$$p(x) = \sum_{y} J(x, y) \qquad \text{Eq. 7(a)}$$

$$q(y) = \sum_{x} J(x, y) \qquad \text{Eq. 7(b)}$$

This makes sense because for *p(x)*, we are not interested in values of *y*, but in the probabilities of *x* regardless of *y*, so for a given *x* we sum over all values of *y* to obtain the marginal distribution *p(x).* A similar argument holds for *q(y)*.

Two random variables are considered "independent" if knowledge of the value of one does not affect our knowledge of the other. Two random variables are independent if and only if their joint probability distribution is equal to the product of their marginal probability distributions, in other words, if

$$J(x,y) = p(x)\, q(y) \qquad \text{Eq. 8}$$

With these concepts in hand, I can now turn to the concept of mutual information. Given two random variables *X* and *Y* defined on the cross product space $x \otimes y$ with joint distribution $J(x,y)$ and marginal distributions $p(x)$ and $q(y)$, mutual information is defined as the relative entropy (KLD or "distance") between the joint distribution function and the distribution that would obtain in the case of complete independence, namely, the product distribution function of the marginals. In the following formula, *MI(X,Y)* is the mutual information. Note that it is symmetrical with respect to *X* and *Y*, that is, *MI(X,Y)* = *MI(Y,X)*:

$$MI(X,Y) = \sum_{x \otimes y} J(x,y) \log_2(J(x,y)\, / p(x)q(y)) \qquad \text{Eq. 9}$$

$$= \sum_x \sum_y J(x,y) \log_2(J(x,y)/p(x)q(y))$$

This definition makes sense if mutual information is considered to be the "distance" between the actual joint distribution and the product distribution. These distributions would be equal (and hence, the mutual information would evaluate to zero) only in the case of complete independence; the further from independence the actual joint distribution is, the more one distribution tells us about the other. Intuitively, mutual information "is the reduction in the uncertainty of one variable due to the knowledge of the other" (ibid., 19).

## 2 Experimental Methods and Results

While experiments have been used to compare many computed measures of rhythmic complexity to perceived rhythmic complexity, to my knowledge, experiments directly involving entropy and rhythm have been rare: an article by Thul and Toussaint (2008), and two by De Fleurian et al. (2014 and 2017). The first of these uses inter-onset intervals to calculate entropy and finds that "[t]he complexity measures based on statistical properties of the inter-onset interval histograms [including entropy] are poor predictors of syncopation or human performance complexity" (663). We will return to these conclusions later.

De Fleurian et al., in both of their papers, ask subjects to listen to test rhythms, to decide whether they should be followed by a note or a rest, and to judge how easy it was to come to a decision. They average the results and correlate them with five computed information theoretic quantities: Shannon entropy, entropy rate, excess entropy, transient information, and Kolmogorov complexity, and find that only entropy rate and Kolmogorov complexity predict experimental findings well.

The present work takes a direct approach - also used by by Thul and Toussaint (2008) - asking subjects to rank eighteen short rhythmic excerpts (seventeen with one duplicate) for complexity.[4] It explores the effects of several variables, based on the hypothesis that they will prove to be significant factors in the perception of rhythmic

[4] This was inadvertent, the result of translating an experiment meant to be conducted in person to a virtual platform. Fortunately, the built-in redundancy of the experiment meant that results were not jeopardized.

complexity: entropy (calculated from IOIs between accented notes in jazz solos), periodicity, syncopation, number/density of notes, and jazz experience level. It takes into account order effects, and, finally, it explores how entropy and the other variables compare and contrast with one another.

It will be important in what follows to distinguish between these related numerical concepts. Note that while in this section, perceived complexity is the dependent variable of interest, in the following section, entropy is the dependent variable of interest. Note that *all* notes in each experimental excerpt are counted in this section, while in the following section using transcriptions, only *accented* notes are counted. Experimental excerpts were constructed from transcriptions by isolating accented notes.

I will adopt a sensitive statistical approach that will allow me to determine the relative role of each of these factors.

N.B.: A statistical confidence level of 95% was used throughout this study.

## 2.1 Experiment Design

Fifteen music majors at UMass Amherst were asked to listen to eighteen short eighth-note-based rhythmic excerpts (seventeen with one duplicate) and to rate them for “complexity” on an integer scale from 1 to 7. A strict definition of “complexity” was not given; subjects were relied upon to provide their own, intuitive, definition. They were asked to listen to each excerpt twice before making a judgement.

The excerpts used for this experiment were derived by selecting excerpts from solos by Louis Armstrong, Coleman Hawkins, Lester Young, Charlie Christian, and Charlie Parker, and isolating only dynamically accented notes.

The excerpts used in this experiment consisted of 20 measures of 4/4 time. In addition to the rhythms of interest, a quarter-note click track emphasizing the first beat of every measure was used. The click track ran for all 20 measures, while most rhythms of interest began after four pickup measures. Some rhythms began after just two or three pickup measures, while some began after five. Most rhythms ended in the 20th measure, but some ended earlier. An example stimulus excerpt is shown in Figure 7.


**Figure 7.** Example of stimulus excerpt

Only notes and rests evenly divisible by eighth-notes were used. The reason for this was that including other rhythms, such as triplet- or sixteenth-note-based rhythms, or rhythms involving ornamental straight-eighth-notes, would have required a longer experiment run time, which was already about twenty minutes. Furthermore, as will be described below, restricting the rhythms to eight-note multiples enables us to quantify syncopation more simply. It will be assumed that the conclusions obtained here, using excerpts divisible by eighth-notes, are generalizable to a broader range of rhythms.

Excerpts were generated using Finale. The woodblock sound on E5 was used for the rhythm of interest, and the clave sound was used to generate the quarter-note click track. Both of these sounds have sharp percussive attacks and no sustain, so that, for example, an eight-note followed by an eighth rest sounded exactly the same as a quarter-note. Excerpts were played at 150 bpm, and swing eighth-notes were generated using quarter-note triplet-eighth-note triplet pairs.

Entropy was the primary variable of interest in this study. In order to study the relationship of entropy to perceived complexity, however, it was necessary to consider other predictor variables as well. First and foremost, it was necessary to control for number of notes.

Since the excerpts were roughly the same number of measures ($\mu = 16.06$, $\sigma = 1.21$, min = 13, max = 18), controlling for number of notes was similar to controlling for *density* of notes; it is not possible to decide *a priori* which quantity a listener would perceive. The importance of this dichotomy will be returned to in Section 3. The variation in excerpt length was the consequence of selecting excerpts from a collection of solos and selecting mainly 16-bar excerpts with one or possibly two pickup measures and sometimes ending early.

Three ranges of entropy were crossed with three ranges of number (or density) of notes and for each of eight combinations, two excerpts were selected. One excerpt, however (high entropy, low number of notes), was also used for the combination high entropy, medium number of notes. Table 2 shows the ranges of entropy and number of notes.

| **Variable/Ranges** | *Low* | *Medium* | *High* |
|---|---|---|---|
| *Entropy* | 1.7 – 2.3 | 2.3 – 2.9 | 2.9 – 3.5 |
| *Number of Notes* | 16 – 23 | 24 – 30 | 31 – 38 |

**Table 2.** Ranges of entropy and number of notes used for selecting excerpts

Note that for computational purposes, variables were treated as continuous; the ranges referred to here were used only for selecting excerpts. For some excerpts, entropy values were on the borderline between two ranges; this did not present a serious problem, since analysis used actual entropy values rather than entropy range.

Next it was necessary to control for the interaction between rhythm and meter. This was done in two ways: by including periodicity as a quantity of interest, and by treating syncopation as a quantity of interest.

To control for *periodicity* of the test rhythms — or lack thereof -– I employed two autocorrelation-like variables I will call *corr.4* and *corr.8*. They represent the number of inter-onset intervals (using the term loosely because they may not be between notes that are adjacent) of duration four or eight eighth-notes, normalized by number of notes. If there is a good deal of correlation at a distance of four or eight eighth notes – in other words, if the test rhythms are strongly tied to the underlying meter – these variables will indicate so.

The formulas are presented below, where *s* is the "signal" (0 or 1) corresponding to a particular excerpt, indexed by eighth-note position in an excerpt *n* eighth-notes long[5]:

$$corr.4 = \sum_{i=5}^{n} s(i)s(i-4)/n \quad \text{Eq. 10(a)}$$

$$corr.8 = \sum_{i=9}^{n} s(i)s(i-8)/n \quad \text{Eq. 10(b)}$$

[5] Results were also calculated using $n^2$ in the denominator instead of *n*. They were essentially unchanged.

Each of these sums reflect the overlap between the signal and a shifted version of the signal. Note that if we shift the signal by four eighth-notes, we can only evaluate the sum from the fifth member of the signal to its end, because the $s(i-4)$ term precludes using anything before the fifth member of the signal. The last term in the sum will be $s(n)*s(n-4)$. No further terms enter into the sum because the signal has only $n$ values. A similar argument holds for *corr.8*.

It is worth noting several things about *corr.4* and *corr.8*.

First and most importantly, while entropy does not depend directly on meter, it depends indirectly on meter via *corr.4* and *corr.8*, since the prevalence of a given IOI – to be specific, an IOI of four or eight – lowers the entropy. The converse is not necessarily true: low entropy may indicate the prevalence of some particular IOI or IOIs, but they need not have values of four or eight. Of course, in this style of music, IOIs of four or eight are bound to be more common than any other IOIs, so in that limited sense the converse *is* true.

Second, while two notes separated by four or eight eighth-notes may add to the sums in Eq. 4.2 (a) or (b) even if there are notes in between them, the normalizing factor $n$ means that pairs of notes separated by four or eight eighth-notes *without* intervening notes contribute slightly more strongly to the sum.

Finally, for low entropy excerpts, the most prevalent placement of notes separated by four eighth notes is on the strong beats one or three, without intervening notes. So for low-entropy excerpts, it is the prevalence of beats one and three that leads to high *corr.4* or *corr.8*, but low entropy.

I also evaluated the syncopation within each excerpt. To control for syncopation, I employed the metric of Longuet-Higgins and Lee (1984), described in Appendix A, divided by number of notes, to yield a variable I will call "LHL quotient" or "*LHLQ*." This variable, by definition, is strongly dependent on meter.

While Smith and Honing (2006) and Fitch and Rosenfeld (2007) have demonstrated the LHL metric to be an empirically meaningful quantity, the fact that a single sounded note can participate in multiple syncopations would seem to be counter-intuitive (see Appendix A), particularly when a sounded note is followed by one or more measures of rest. In such a situation, a single note – say on the and of four – can be followed by syncopation totals of 7 (a large number in this context)[6] for each measure of rest following the sounded note. In the excerpts used here, there are frequently one or two measures of rest, while in the excerpts used by Smith and Honing (who use the data of Shmulevich and Povel, 2000) and by Fitch and Rosenfeld, there do not appear to be any measures of rest. This should be addressed in a later study.

Next, order effects were taken into account. (Note that eight different excerpt orderings were used). It was found that a relatively high complexity rating exerted an upward pressure on the following rating, while a relatively low complexity rating exerted a downward pressure on the following rating. This effect was significant in all models. It was, however, a small effect, and did not qualitatively affect the outcomes of the experiment.

One final factor was taken into account: the jazz experience level of the participants. This was recorded as a binary variable indicating whether or not the subject

[6] This assumes that only eighth-note subdivisions are used, not sixteenth-note subdivisions. Including sixteenth-notes would not assuage the problem; in fact, it would make it worse.

was a jazz musician (undergraduate music majors at UMass Amherst study either jazz or classical music). It did not appear to make any difference in the analysis, and was therefore ignored.

## 2.2 Experimental Results

Data were analyzed using a mixed effects multivariate regression model[7]; the Satterthwaite approximation for effective degrees of freedom was used since the variance of the data was not known *a priori*. A "mixed effects" model was necessary because some factors were "fixed effects" (entropy, periodicity, syncopation, carryover effects, number/density of notes), while others were "random effects" (random intercept terms for subject and for excerpt). In general, a fixed effect is one whose specific values we are interested in, while a random effect is one that is chosen merely to be a representative of a larger population of values. Data were analyzed treating all factors of interest as continuous variables (this does not include experience level), and also treating the perceived complexity ratings as a continuous variable.[8]

Four conditions for multiple regression were tested: independence, normality, homoschedasticity, and linearity. While the independence condition is not met since the model includes carryover effects (the influence of the previous complexity rating on the current complexity rating), this was judged not to be a problem by statisticians Anna Liu and Michael Lavine.

---

[7]. On the advice of Anna Liu, personal communication

[8] The question of whether or not Likert–scale variables – such as the response variable in this experiment – can be treated as continuous variables is hotly debated, as a quick google search will indicate. There seems to be a consensus that a minimum of seven response choices (such as used in this experiment) is required for such a variable to be treated as continuous. However, personal communications from Anna Liu indicate that it is probably fine to treat the Likert–scale data as continuous for this experiment. This is a common assumption in the literature, and it is the assumption I adopt here.

As a first step in analyzing the data, a backwards selection algorithm was used. This algorithm starts with a full model, and progresses in a step-wise fashion by deleting one variable at a time according to which change will yield the greatest reduction in the Akaiki Information Criterion until some prespecified stopping condition is reached (all of this was done using the 'step' command in the programming language R). The backward selection algorithm does not necessarily yield the absolute *best* model, but is merely a heuristic to identify a *plausible* model for further consideration.

The Akaiki Information Criterion, and its relative the Bayesian Information Criterion, are both derived from the log likelihood function. The likelihood function treats observed variables as parameters and fit parameters as variables, and is maximized for a maximum-likelihood fit. Using the logarithm of the likelihood function is more convenient than using the likelihood function itself due to the very small numbers involved.

Both the AIC and the BIC yield smaller values for better models. Both penalize a large number of parameters and favor simpler models. The AIC tends to be more liberal in terms of including parameters, while the BIC is more conservative. The Akaiki Information Criterion is defined by:

$$\mathrm{AIC} = 2k - 2ln\hat{L} \qquad \text{Eq. 11}$$

where $k$ is the number of parameters in the model and $\hat{L}$ is the maximized likelihood function. The Bayesian Information Criterion is defined by:

$$\mathrm{BIC} = k\, ln(n) - 2ln\hat{L} \qquad \text{Eq. 12}$$

where $n$ is the sample size.

The typical mixed-effects multiple regression fit in the statistics programming language R uses the Restricted Maximum Likelihood (REML) of a model for a set of data. However, fitting data this way precludes the use of AIC to compare models exactly, including in the process of backwards selection, since the AIC is based upon maximizing the likelihood function rather than the "restricted" maximum likelihood. Furthermore, it complicates the process of testing for interaction effects. Therefore, I opt for a "maximum likelihood" method rather than the "restricted maximum likelihood" method here during model identification and interaction testing, then switch back to REML for reporting the final results. Choosing REML methods over maximum-likelihood methods in the final stage of calculations does not qualitatively affect final *p*-values.

With this in mind, I return to the backward selection algorithm. Starting with a model including all of the fixed effects and random effects listed above, it identifies as a good model one that includes: *corr.4*, *LHLQ*, number of notes, and order effects as fixed effects, and random intercept terms for subject and excerpt. The coefficient corresponding to *corr.4* is negative, indicating that perceived complexity has an inverse relationship to *corr.4*: decreasing periodicity increases complexity rating. The other coefficients are positive. This model will be called the long model.

Since I am primarily interested in entropy, I also propose a model using entropy, number of notes, and order effects as fixed effects, and random intercept terms for subject and excerpt. This will be called the short model.

Both the short model and the long model were tested for interactions between fixed effects. Interactions occur when the effect of one variable depends on the value of another. There are multiple ways of testing for interactions; the method I used employed the ANOVA procedure to compare two models at a time, one being the model without interactions and the second cycling through all possible combinations of fixed effect pairs, triples, and (for the long model) quadruples. No significant interactions were found.

Table 3 shows the AIC and BIC for the long model and the short model, as well as the $R^2$ metric for both models (the Nakagawa $R^2$ for mixed-effects models was used); lower values of the AIC and BIC indicate a better fit, while higher values of $R^2$ indicate a better fit. The rule of thumb for the AIC and BIC is that a difference of 2.0 or less is insignificant, while differences of 10.0 or more are extreme. As Table 6 shows, the long model is better than the short model according to the AIC and BIC, while the short model is better than the long model according to the $R^2$ metric. $R^2$ values in the 0.5–0.6 range indicate robust models, particularly for studies involving human psychology and performance (online consensus).

| **Model/Criterion** | **AIC** | **BIC** | $\boldsymbol{R^2}$ |
|---|---|---|---|
| *Long Model* | 714.14 | 742.40 | 0.514 |
| *Short Model* | 722.31 | 747.04 | 0.527 |

**Table 3.** Goodness of fit comparisons using Akaike Information Criterion, Bayes Information Criterion, and $R^2$

*P*-values for the long model are shown in Table 4, and reveal that all included fixed effects are significant. This is not surprising, as we expect periodicity, syncopation, number of notes, and order effects to be reflected in complexity ratings. The *p*-values for the short model are shown in Table 5, and reveal that all included fixed effects are significant. This, too, is to be expected, as we expect entropy, number of notes, and order effects to be significant. The conclusion is that the models based on lack of periodicity/ syncopation and on entropy are both interesting and informative models.

| **Variable** | ***p Value*** |
|---|---|
| Number of notes | $2.83*10^{-5}$ |
| *Corr.4* | $1.20*10^{-5}$ |
| *LHLQ* | 0.0154 |
| Order effect | $2.14*10^{-5}$ |

**Table 4.** Description of long model

| **Variable** | ***p Value*** |
|---|---|
| Number of notes | 0.000152 |
| Entropy | 0.010499 |
| Order effect | $3.62*10^{-5}$ |

**Table 5.** Description of short model

An issue that can adversely affect the interpretation of multiple regression results is “collinearity” or “multicollinearity”. This is a situation that obtains when one or more predictor variables can accurately be derived from one or more other predictor variables by linear combination. This is undesirable since in this situation, the coefficient estimates of the multiple regression may change erratically in response to small changes in the model or the data (*Wikipedia*). An easy way to test for multicollinearity is to use the so-called *variance inflation factor*. This represents the ratio of the variance as it occurs for each term in a model to what that variance would be in the case of perfect independence between variables. VIFs greater than 2.5, 5.0, or 10.0 – depending on the analyst's choice – reveal the presence of multicollinearity.

Applying this technique to the models at hand, using thresholds as low as 2.5, revealed that multicollinearity is not a problem here. Calculating VIFs for the fixed effects in the "full" model, however (entropy, *corr.4*, *corr.8*, *LHLQ*, carryover effects, and random intercept terms for subject and excerpt) yields the following VIFs (Table 6):

| **Variable** | Entropy | *corr.4* | *corr.8* | *LHLQ* | Number/Density of Notes | Carryover Effects |
|---|---|---|---|---|---|---|
| **VIF** | 7.45 | 7.03 | 3.52 | 2.61 | 2.93 | 1.02 |

**Table 6.** Variance inflation factors (VIFs) for full model

Clearly, entropy and *corr.4* exhibit collinearity, which is not surprising given that they exhibit strong negative correlation ($r$ = -0.88). Note that entropy and *corr.8* have a strong negative correlation, too ($r$ = -0.84) but the V.I.F. for *corr.8* is only 3.51.

**2.3 Discussion**

While it may appear that the models presented above are independent of one another, they may in fact be an example of "statistical mediation." Statistical mediation happens when an independent variable "X" influences a dependent variable "Y" through a mediating variable "M." In the current situation, X could be "entropy," Y could be "perceived rhythmic complexity," and M could be "*corr.4*."

This is supported by the fact that there is a strong negative correlation between entropy and *corr.4* (–0.88), which is stronger than the correlation between entropy and *corr.8* (–0.84) or entropy and syncopation (0.35). It is also supported by the fact that the variance inflation factors for entropy and *corr.4* are high (about 7.0) while those for *corr.8* and syncopation are much lower (<3.5).

Furthermore, strong evidence for statistical mediation comes from the following observation, found on the University of Virginia's website:

*If a mediation effect exists, the effect of X on Y will disappear (or at least weaken) when M is included in the regression. The effect of X on Y goes through M. If the effect of X on Y completely disappears, M fully mediates between X and Y (full mediation).*

www.library.virginia.edu/data/articles/introduction-to-mediation-analysis

Starting with the short model (entropy, num. notes, order effects, and random intercept terms for subject and excerpt), the *p* value for entropy is 0.000152. Adding *corr.4* completely obscures this effect ($p$ = 0.30764), showing that *corr.4* completely mediates between entropy and perceived rhythmic complexity.

Adding *corr.8* instead of *corr.4* results in a *p* value for entropy of 0.06013, thus demonstrating that *corr.8* partially mediates between entropy and perceived complexity. Adding *LHLQ*, on the other hand, does not obscure the effect of entropy at all: $p$ = 0.000436.

Finally, it is possible that mediation effects occur between *corr.4* and *corr.8*. This is of secondary interest, however, and will not be pursued further.

What all of this means is that for excerpts with high values of *corr.4* or *corr.8* there are many IOIs of four or eight eighth-notes, and thus a low entropy, while for excerpts with a low value of *corr.4* or *corr.8*, there is a more diverse distribution of IOIs and thus a higher entropy. So it appears that this is a case of statistical mediation, which does not change the fact that entropy alone significantly affects complexity ratings.

Overall, then, the experiment revealed that for rhythms comprised solely of eighth-notes and rests and their integer multiples, entropy differences are reflected in subjective rhythmic complexity ratings (though this is probably mediated by periodicity); that periodicity or the lack thereof and syncopation are reflected in complexity ratings; that the number of notes (when excerpt length is held roughly constant) is reflected in complexity ratings; and that carryover effects are reflected in complexity ratings. More work on the effects of jazz experience level may be called for.

These results challenge the finding of Thul and Toussaint (2008) that findings based on the statistical properties of IOI histograms (e.g. entropy) do not predict human performance metrics. There could be several reasons for this. Perhaps the null result regarding human performance metrics reflects a fundamental difference between measuring human performance metrics and measuring subjective complexity ratings. Perhaps the null result discrepancy comes from the fact that Thul and Toussaint used sixteenth-note subdivisions in addition to eight-note subdivisions; this methodology may or may not be more accurate than using just eighth-note subdivisions. Or perhaps these are merely results that disagree with each other, calling for more work to determine which is correct. In any case, further studies are called for.

In a future study, it would be interesting to study the carryover effects of actual independent variables (entropy, *corr.4*, *corr.8*, *LHLQ*, number/density of notes,). In this scenario one might expect a high entropy value, for example, to exert a downward pressure on the following rating, while a low entropy value might exert an upward pressure on the following rating. In other words, having just heard a highly entropic excerpt, for example, a subject might hear the following excerpt as being *less* entropic by comparison.

## 2.4 Conclusion

In this experiment, fifteen music majors rated eighteenth rhythmic excerpts (seventeen with one duplicate) for complexity. Two models were used to describe the data well ($R^2 \approx 0.52$): one including entropy (calculated from inter-onset intervals between notes in each excerpt), number of notes, carryover effects, and random intercepts for subject and excerpt; and one including periodicity, syncopation, number of notes, carryover effects, and random intercepts for subject and excerpt. It is likely that the two models demonstrate statistical mediation, in which entropy influences periodicity (or lack thereof), and periodicity in turn influences perceived complexity.

In this experiment, only rhythms divisible by eighth-notes were used. This was necessary because including a wider array of rhythmic subdivisions would have made the experiment's run time too long. Another experiment including a wider array of rhythmic subdivisions will have to await another study. The results presented here, however, already suggest strongly that perceived rhythmic complexity depends on entropy, when entropy is calculated using probability distributions obtained from measuring IOI's between dynamically accented notes in jazz solos. In the following chapters, it will be assumed that entropy is, indeed, perceptible. If it is not, however, it is interesting in its own right, as a signature of musical style. This will be demonstrated in the next section.

## 3 Computational Results

My main goal in this section is to understand whether entropy, on the whole, is a signature of musical style among the five musicians considered in this study. This is an interesting question because the results from Section 2 seem to indicate that entropy is an indicator of rhythmic complexity, though the transcribed solos analyzed here frequently contained a wider range of rhythmic subdivisions than those contained in the experimental excerpts. To be specific, they contained eighth-note and quarter-note triplets, sixteenth-notes, and ornamental straight-eighth-notes. Even if entropy is *not* an indicator of rhythmic complexity for this wider range of rhythmic subdivisions, however, the question of whether or not entropy is a signature of style is still an interesting one, if for no other reason than academic curiosity.

### 3.1 Corpus of Transcribed Solos

The computations described here are based upon a corpus of 88 transcriptions of solos by five great jazz musicians: Louis Armstrong (1901–1971), Coleman Hawkins (1904–1969), Lester Young (1909–1959), Charlie Christian (1916–1942), and Charlie Parker (1920–1955). Note that the birthdates of these musicians span the first two decades of the twentieth century. The transcriptions were done by the author, though the *Omnibook* was used as a starting point for the Charlie Parker transcriptions.

Transcribed solos were converted to Excel files by evaluating the elapsed times, in eighth notes, from the beginnings of solos to the onset of accented notes. For some double-time solos, e.g. on ballads, elapsed times were sometimes evaluated in sixteenth notes. Eighth-notes (or sixteenth-notes on some double-time solos) were assumed to be swung unless they were flagged as "straight."

Efforts were made to select solos representative of each soloist's different style periods, and to include chronologically even representations of each soloist's oeuvre. For Armstrong, solos were included from his Hot Fives and Hot Sevens, from Louis Armstrong and His Savoy Ballroom Five, from Louis Armstrong and His Orchestra, and from Louis Armstrong and His All-Stars. (His early work with King Oliver was not included). Representatives of Armstrong's trumpet solos and of his vocal work were included. For Hawkins, his work with Fletcher Henderson was represented, as was his famous recording of "Body and Soul" from 1939 (his work in Europe was not included), his tenure as a leader on 52$^{nd}$ St., his work with Thelonious Monk, and his later work with pianists Tommy Flanagan and Paul Bley. Lester Young's work with Basie was represented, as was his work with Billie Holiday, his work as a leader both before and after his enlistment in the Army, and his late work with Jazz at the Philharmonic. Charlie Christian's premature death resulted in a lack of discernible style periods; the years 1939–41 were covered roughly equally. Finally, the solos of Parker, according to Kernfeld (1996), can be divided into his early style (pre-1944) and his mature style (post-1943). Both periods are represented here.

All excerpts were in 4/4 time, sometimes articulated as 12/8 time, with a tempo range of 63 (Armstrong "What a Wonderful world") to 280 bpm (Parker "Honeysuckle Rose" and "Crazeology").

The smallest subdivision used in the transcriptions was the sixteenth note (or in the case of double time solos, the thirty-second note). This subdivision was sufficient for identifying accents in the solos transcribed. Efforts were made to accurately represent rhythmic anticipations and suspensions that approximate eighth-note rhythms, for example in Armstrong's solos on "Stardust" or "What a Wonderful World." Straight eighths were distinguished from swing eighths, and in some solos based on double time (e.g. Hawkins, "Wanderlust"), sixteenth notes were treated as swung notes, with some straight sixteenths.

The shortest excerpt was 12 bars with pick-ups (Parker, "Hootie Blues"), while the longest was 200 bars (Young, "Ad Lib Blues"). The question of minimum solo length is an important one, since as the number of bars tends toward zero, so must the excerpt's entropy. For the time being, 12 bars were considered long enough to produce a meaningful entropy value. As will be discussed later, however, solo length will be accounted for by treating the number of accents as a covariate in the analysis of entropy's dependence on musician.

Only dynamically accented notes were used in the calculations. Accented notes were used because they reflect a layer of structure superimposed by the soloist on the structure of the solo. Dynamically accented notes were used on the assumption that they are the easiest to perceive. In a future study, other kinds of accent might be included.

Joel Lester, quoted in John Roeder's "A Calculus of Accent," defines accent in terms of the music surrounding a given accent, specifically "the relative strength of a note or other musical event in relation to surrounding notes or events." An archetypical series of accents, from Lester Young's solo on "Blues For Greasy," is shown in Figure 8 (this was seen earlier in Figure 1). A more complex example, from Coleman Hawkins's solo on "Body and Soul," is shown in Figure 9; here a phrase played *mezzo piano* is followed by a phrase played *mezzo forte* containing an accent on the second beat of the second phrase. The first note of the second phrase is *not* considered to be an accent, even though it is louder than the preceding phrase. Two examples from the Louis Armstrong trumpet corpus illustrate how accents interact with register; this is of interest because of the strong dependence of loudness on register in trumpet music. In the first example, from his solo on *Stardust*, an ascending phrase lands on an accented note which is then repeated with accents (Figure 10). I perceive these notes as being accented even though there is a crescendo leading up to them. An example from "I Double Dare You," by contrast, is in a moderately high register, but does not accent the highest note in the phrase (Figure 11). Finally, I discuss one of Armstrong's vocal solos on "What a Wonderful World" (see Figure 17, later in the document). On the bridge, there are many repeated accented notes, but there are enough notes in the solo that are *not* accented to make the identification of accented notes meaningful.

Most of the eighth note pairs considered here are swing eighths, meaning that they can be grouped into long-short pairs starting on quarter-note beats. According to Collier and Collier (2002), for example, Armstrong sometimes employs eighth note pairs with a durational ratio of 1.6:1 (or 8:5). In many excerpts, however, Armstrong and other soloists employ a ratio of 2:1.


**Figure 8.** A straightforward example of accents from Young, "Blues For Greasy"

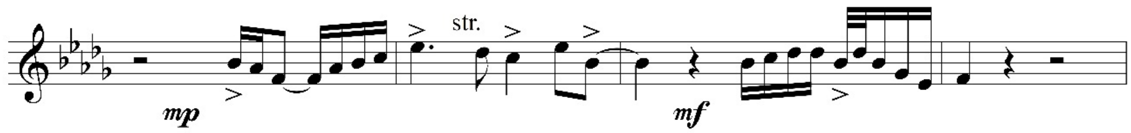


**Figure 9** An excerpt from Hawkins's "Body and Soul" in which a change in dynamic levels does not cause an accent

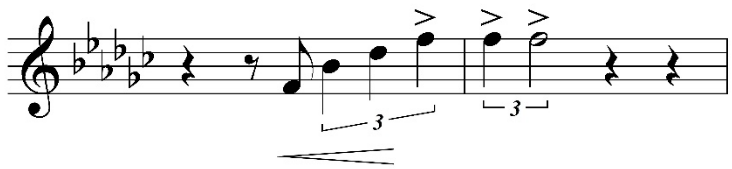

**Figure 10.** An excerpt from Armstrong's "Stardust" in which a crescendo does not negate the perception of accents

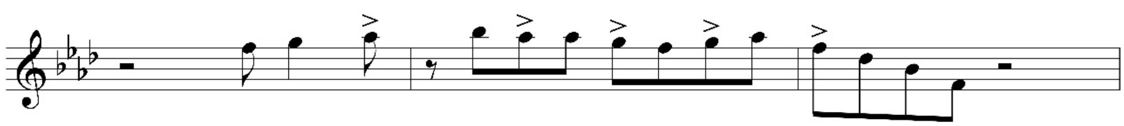

**Figure 11.** An excerpt from Armstrong's "I Double Dare You" in which register does not solely determine accents

A ratio of 2:1 creates a triplet feel, since it corresponds to a single quarter-note triplet followed by a single eighth-note triplet. This ratio implies a compound 12/8 meter, and the rhythm section often supports this; a prime example is one of Armstrong's performances of "What a Wonderful World," from 1967, transcribed in Figure 17. The piano plays chords in 12/8 time, and this provides the backbone upon which Armstrong's solo is built.

For some excerpts, it must be specified whether the swing ratio is 2:1 or not in order to calculate entropy from the transcribed solos, since if it *is*, the IOIs involving swing eighth notes are equivalent to rhythms involving triplets, thus potentially reducing the total number of distinct IOIs in the excerpt. This causes a potential problem since there is theoretically a discontinuity in entropy as a function of swing ratio: as 2:1 is approached, the entropy is calculated according to the formula for *non*-triplet swing eighth notes, but this jumps to the *triplet* version of the entropy when 2:1 is reached. For other excerpts, this does not matter because there are no accents on the third triplet of any quarter-note beat.

For ratios other than 2:1, the exact ratio is unimportant; what matters is that the difference between groups of eighth note beats starting on a "long" eighth note rather than a "short" eighth note is taken into account in the analysis. This is true for groups consisting of odd numbers of eighth notes. For example, a group of three eighth notes starting on a long eighth note beat would have a duration of L+S+L, while the duration of three eight notes starting on a short eighth note beat would be S+L+S, where S and L stand for the (unequal) durations of the short and long eighth note beats respectively.

To facilitate analysis, only excerpts in which the swing ratio is obviously 2:1 or in which the entropy does not depend on the swing ratio were included in the corpus. This excludes excerpts in which the swing ratio is difficult to ascertain *and* in which the entropy depends on whether or not the swing ratio is 2:1.

One solo had to be rejected because it was constructed almost entirely of straight or almost straight eighth notes: Charlie Parker's solo on "KoKo." Solos such as these cannot be included in the corpus, because the decision of whether to treat them as swing or straight in calculating entropy has a marked effect on the outcome (on the order of 10% for "KoKo").

For a chart containing musician names, excerpt names, recording dates, a binary variable called "TS" for "triplet swing," which indicates whether or not an excerpt exhibited a 2:1 swing ratio AND the entropy depended on whether or not the swing ratio was treated as 2:1, number of distinct IOIs, number of accents, and entropy, see Appendix B at the end of this paper.

### 3.2 Methods

The primary tool employed for the task of determining whether or not entropy depends on musician is that of *estimated marginal means*, or EMMs. EMMs allow us to study the dependence of a continuous "response" variable (in this case, entropy), on a single discrete "factor" (in this case, musician), in the presence of continuous "covariates" which may depend on the factor. *P* values for each pairwise combination of musicians reveal whether or not entropy depends on musician. EMMs are calculated from multiple linear regression models; four conditions for regression were verified: independence, normality, homoschedasticity, and linearity.

As covariates to include in this model, I selected number of distinct IOIs and number of accents. Not only are these intuitive choices – since both of them could be reflected in entropy values and both of them could depend on musician in the corpus used here – they also sidestep the issue of normalization. Rather than normalize by maximum entropy for a given number of IOIs, I use number of IOIs as a covariate. Rather than normalize by sample length, I use number of accents as a covariate. This is also a convenient way to sidestep the issue of sample size (see below).

To compare different musicians in terms of entropy, the present study takes a different approach from that used in previous studies (Youngblood 1958; Knopoff and Hutchinson 1981, 1983; Snyder 1990): rather than grouping data together by composer, different excerpts are evaluated for entropy separately, and the estimated marginal means technique lets us evaluate differences in entropy for significance between musicians, obviating the need for a minimum number of points for each musician (though a large number of points spread out over the excerpts was still used – see below).

In order to justify the inclusion of number of distinct IOI's and number of accents as covariates, I turn to a graphical technique called the "added variable plot." Added variable plots reveal whether or not it makes sense to add an independent variable to a linear model. First, one constructs two models: $y \sim x_1 + x_2 + \ldots + x_n$ and $x_{n+1} \sim x_1 + x_2 + \ldots + x_n$, where $y$ is the dependent variable, $x_1 \ldots x_n$ are the independent variables already included in the model, and $x_{n+1}$ is the independent variable to be tested for addition. One then calculates the residuals of the two models and plots them against each other. A straight line indicates that the variable $x_{n+1}$ should be added to the model.

In this case, I begin by considering the addition of "number of distinct IOIs" to the single variable "musician"; the corresponding plot, Figure 12, shows that the variable *should* be added. The same procedure is used to show that "number of accents" should be added to "musician" and "number of distinct IOIs" (Figure 13). According to statistician Michael Lavine (personal communication), these are "textbook examples."

Having selected number of distinct IOIs and number of accents as covariates based on our intuition about the variables (other than musician) upon which entropy should depend, and based on the added variable plots, I use another graphical technique, the residuals vs. fitted values plot, to test whether or not the model is a good fit to the data, and whether or not more variables should be added. A complete, well-fitting model is indicated by a random scattering of points, showing that there is no systematic error in the data. The residuals vs. fitted values plot for the model entropy ~ musician + number of IOIs + number of accents is shown in Figure 14; it reveals that the model is a good fit to the data, and that adding more variables might result in overfitting.

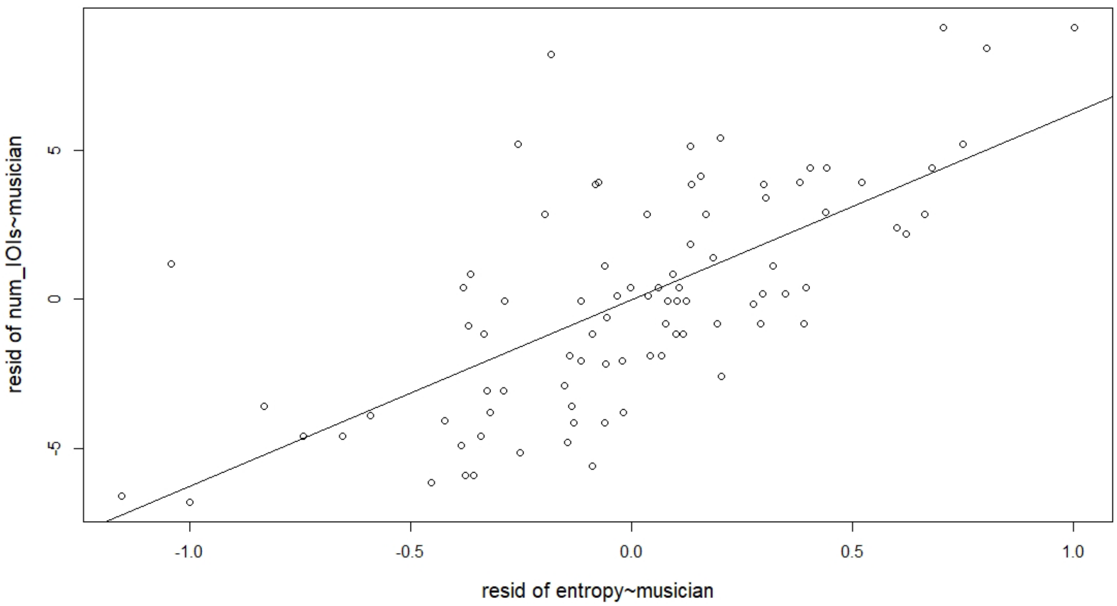


**Figure 12.** Added Variable Plot for number of IOIs

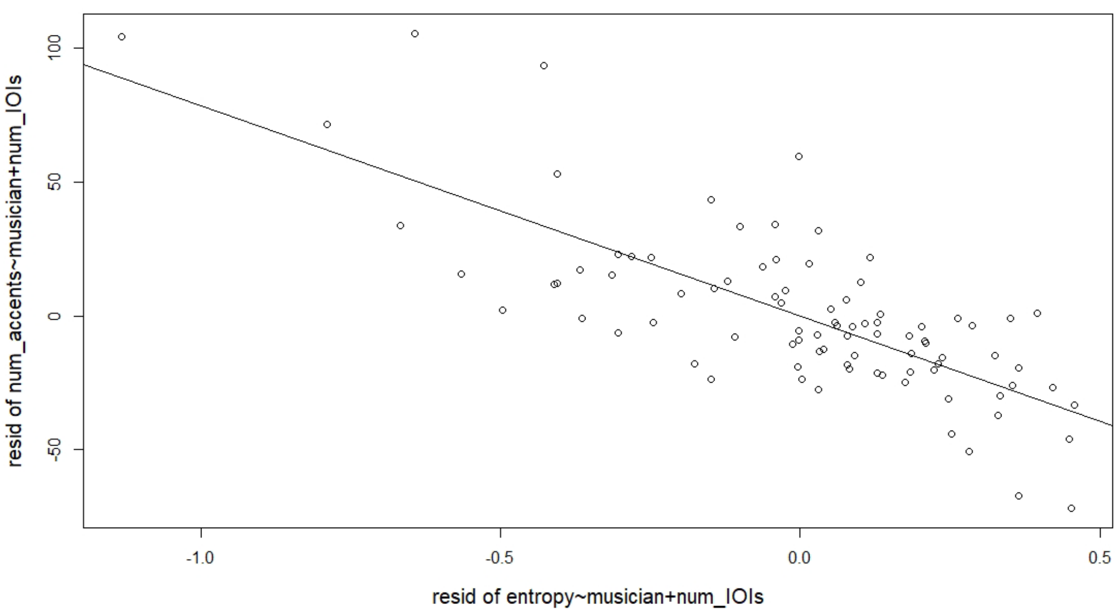


**Figure 13.** Added Variable Plot for number of accents

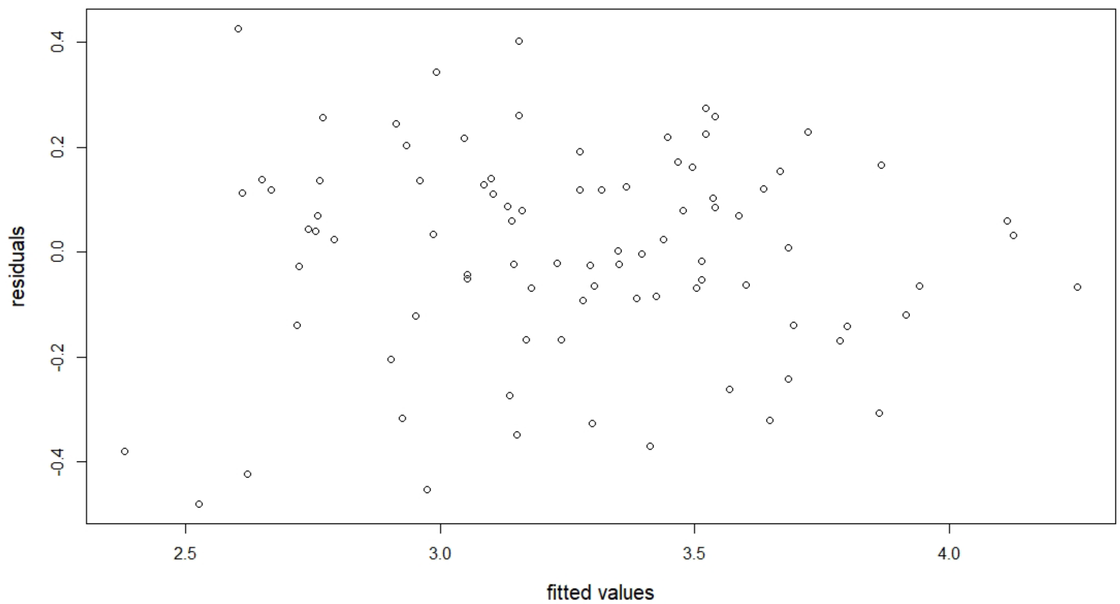


**Figure 14.** Residuals vs. Fitted Values Plot for two-covariate model

.

Recall that in the experimental portion of this paper, I raised the issue of multicollinearity, and found that the variables used there were not collinear. In the present scenario, it is again desirable to verify the lack of collinearity in the data. Here, I use *generalized* variance inflation factors due to the fact that the variable *musician* has four degrees of freedom. After calculating the GVIF's and transforming them using the formula $GVIF^{0.5\nu}$ where ν is the number of degrees of freedom, it can be concluded that there is no collinearity.

In a situation such as this, when comparisons are made between multiple pairwise combinations, *p* values must be adjusted to account for the multiple comparisons. That is because if an individual test has a 0.95 confidence level, several simultaneous tests will have a lower confidence level: increasing the number of tests creates the opportunity for more Type I errors. Increasing the individual *p* values corrects this problem. The simplest and most conservative way to do this is to use the *Bonferroni correction*, which multiplies each pairwise *p* value by the number of pairwise combinations. So in the present scenario, with five musicians, there are 5*4/2 = 10 pairwise combinations, and we must multiply each unadjusted *p* value by 10. (Resultant *p* values greater than one are automatically assigned the value 1.0).

Here I digress for a moment to describe an anomaly in the data. In his solo on *I Never Knew*, Armstrong uncharacteristically plays two "clams," in measures 9 and 10, on beats four and one respectively. To examine the effects of these notes on entropy, I treated them four ways: neither as an accent; the first of the two as an accent; the second of the two as an accent; and both as accents. The net effect on entropy was less than five percent, the effect on num_IOI's was at most about eight percent (one out of twelve or thirteen), and the effect on num_accents was less than six percent (one or two in the range 33–35). None of these differences manifested as appreciable changes in *p* values for the models discussed here, but to be as accurate as possible, *p* values – as well as $R^2$values - were averaged across each of the four possible treatments of the clams[9,10].

One final topic pertaining to the analysis of this data must be addressed, as in the experimental case: interactions. As in the experimental case, this is handled by performing an analysis of variance on pairs of models having no interaction terms or having one or more interaction terms. This revealed no interactions.

### 3.3 Results

[9] Note that the quantity in question is not perceptual, but intentional: what did Louis Armstrong *intend* to play? We can never know, so we average over the four possibilities.
[10] The added variable plots displayed above had almost no perceptible dependence on disposition of the clams.

Table 7 shows the pairwise *p* values resulting from applying the estimated marginal means technique; in each pair, the lower value is listed first. I find that while Armstrong, Christian, Hawkins, and Parker are all less entropic than Young, the difference between Hawkins and Young is not statistically significant since the *p* value is greater than 0.05. In particular, the *p* values for Armstrong/Young, Christian/Young, Hawkins/Young, and Parker/Young are 0.0224, 0.0239, 0.2404, and 0.0011. Pairs not involving Young all have *p* values of 1.0, which occurs whenever multiplying a raw *p* value by the Bonferroni correction results in a number greater than or equal to one.

The adjusted $R^2$ metric for this model was high: 0.7896. This indicates that most of the variation in entropy is due to the independent variables, rather than to noise.

| **Musician Pair** | ***p Value*** |
|---|---|
| Armstrong/Young | 0.0224 |
| Christian/Young | 0.0239 |
| Hawkins/Young | 0.2404 |
| Parker/Young | 0.0011 |

**Table 7.** Pairwise comparisons using the EMM procedure; comparisons not listed have *p* values of 1.0

Care must be taken in evaluating these results to take researcher error into account, since in a manually transcribed corpus of this size (6,299 accented notes; 20,000–30,000 notes total) there are bound to be errors. This is discussed in detail in Appendix C; the upshot is that random error probably does not qualitatively affect the above-reported *p* values, but there is a possibility that the Christian/Young comparisons should be reported as only *marginally* significant ($p \approx 0.05–0.06$), or even possibly insignificant. The pairwise comparisons between Armstrong and Young and Parker and Young, however, are secure.

I can take the analysis one step further by identifying outliers in the data. This is done by fitting the data with a multiple regression model and calculating residuals; points corresponding to residuals that are greater in absolute value than twice the standard deviation of the residuals are considered to be outliers. Using this method, five outliers were identified, including, notably, two early Charlie Parker excerpts, "Moten Swing" and "Honeysuckle Rose."

In light of the above discussion regarding Armstrong's solo on "I Never Knew," the identification of outliers was carried out for all four dispositions of the accented/non–accented clams. Results were consistent for all four treatments of the problematic notes.

| **Musician Pair** | ***p Value*** |
|---|---|
| Armstrong/Young | 0.00035 |
| Christian/Young | 0.0011 |
| Hawkins/Young | 0.2632 |
| Parker/Young | 0.0006 |
| Armstrong/Hawkins | 0.4990 |
| Christian/Hawkins | 0.6673 |

**Table 8.** Pairwise comparisons using the EMM procedure, data with outliers deleted. Pairs not shown have $p$ values of 1.0

Pairwise comparisons using the modified data are shown in Table 8. Removing outliers did not qualitatively change the results, but rather intensified the contrasts between Armstrong and Young, Christian and Young, and Parker and Young. *P* values were 0.0004, 0.0011, and 0.0006, respectively. The difference between Hawkins and Young was still not significant ($p = 0.2632$). All other *p* values were 1.0 with the exception of Armstrong/Hawkins ($p = 0.4990$) and Christian/Hawkins (0.6673). The adjusted $R^2$ value for the data with outliers deleted was 0.818, even higher than that with outliers included. (Note that checking for interactions in the data with outliers excluded indicated that it was not important to include interaction terms).

### 3.4 Debunking Hypotheses About Chronological Trends

Next I explore a way of parsing the data on Armstrong into early and late periods suggested by Barry Kernfeld (1996). Kernfeld posits that Armstrong's work from *before* 1936 is qualitatively different from his work from *after* 1936; he states that after 1936, Armstrong basically pandered to the masses, while his earlier work was more complex. I tested this hypothesis in terms of entropy using the EMM procedure, and found that there was no significant difference between the two periods ($p = 0.221$). Regarding Kernfield's assertion that Armstrong's later work pandered to the masses, one must only listen to his popular recording of "What a Wonderful World" (1967) to see that Armstrong could gain the widespread adulation of fans without sacrificing musical complexity (entropy of 3.877).

Again using the EMM technique, I explore the commonly held belief that Lester Young's playing was altered by his stint in the Army, and find that entropy does not reveal a difference between his pre- and post-Army solos ($p = 0.229$).

### 3.5 Hawkins vs. Parker

The fact that the pairwise comparison between Hawkins and Parker ranks Hawkins above Parker (though not significantly) merits some attention, given that Parker is usually considered to be more rhythmically complex – and presumably more entropic – than Hawkins.

For example, according to Scott DeVeaux in The Birth of Bebop, Hawkins was known for his rhythmic predictability:

*His notorious tendency toward rhythmic uniformity – a steady stream of eighth notes … later caricatured by one writer as a "machine-gun style" – was only exacerbated by his ongoing project of crowding as much of the underlying harmony as possible into his improvised line.*

DeVeaux 1997, 85

Martin Williams confirms this observation:

*Rhythmically, [Hawkins] continued to live in the early 'thirties – but, again, with more regular accents than many players of that period.*

Williams 1993, 77

Parker, on the other hand, was known for his rhythmic adventurousness:

*… the pattern of accents in a Charlie Parker line is in a constant state of flux – falling sometimes on the strong beats of the measure, but also (and quite unpredictably) on "weak beats" (beats 2 and 4) or on the weak half of the eighth note pair.*

DeVeaux 1997, 264

If entropy does not reflect the predictability of Hawkins and unpredictability of Parker we are led to expect from conventional wisdom, what accounts for these expectations? One possibility is that Hawkins uses more dynamical accents than Parker, perhaps creating an expectation of accents. This theory is supported by the fact that the average dynamical accent density (number of dynamical accents divided by excerpt length) for a limited corpus of thirteen Hawkins excerpts is 1.74 while for a corpus of fifteen Parker excerpts it is 1.36[11]. A two-sample *t*-test yields a *p* value of 0.0046; so the difference is statistically significant.

I can go one step further, and obtain results based not only on *dynamical* accents but on *contour* accents as well[12]. Contour accents occur when the melody changes direction; only changes in direction prepared by two intervals moving in the same direction were counted. This is a reasonable thing to do, given that, to my ears, Parker's lines owe much of their interest to changes in direction. For example, the lick shown in Figure 15 changes direction eight times in the space of just two bars. This lick appears in "Au Privave," "Billie's Bounce," and "Now's The Time." (To be fair, this lick also appears once in a Hawkins solo, but not frequently as in the Parker corpus). Using the same samples as those used for the dynamical accent calculations[13], I find that the average contour accent density for Hawkins is 1.03, while for Parker it is 1.34. Again using a two-sample *t*-test, I find that $p = 0.00078$. Thus, there is a significant difference in contour accent density between Parker and Hawkins, perhaps creating the expectation of unpredictability, unlike the expectation created by greater density of dynamical accents.

---

[11] The distribution of dynamical accents for a corpus of sixteen Parker excerpts was found not to be normal, a requirement for using the two-sample *t* test used here. Deleting a single excerpt ("Hootie Blues") resulted in a normal distribution.

[12]For more on contour theory, see Deutsch (1972), Dowling (1978), Edworthy(1985), Friedmann (1985), Polansky (1996), Quinn (1999), and Schmuckler (1999).

[13]Including or excluding "Hootie's Blues" did not affect the normality of the contour accent sample, so it was included in the two-sample *t* test.

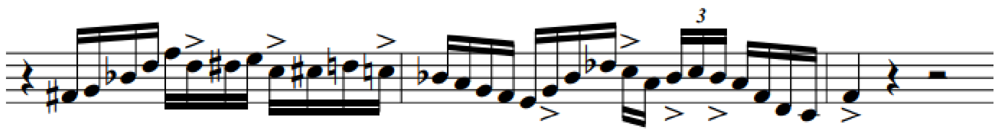

**Figure 15.** Charlie Parker lick changing direction eight times in two bars

### 3.6 Number/Density of Notes

Recall from Section 2 that entropy, to the extent that we can assume it affects the perception of rhythmic complexity for rhythms that are not eighth-note-based, is not the only factor affecting the perception of rhythmic complexity. Periodicity (or lack thereof), syncopation, and number of notes (for an experiment involving approximately equal excerpt lengths), for eighth-note-based rhythms, are all factors of interest. Recall that in Section 2 I pointed out that it cannot be decided *a priori* whether listeners respond to note number or note density in formulating complexity ratings.

Note that in designing the estimated marginal means model for entropy, accent number was used as a covariate rather than accent density because I desired a variable that reflected length of excerpt and because the added variable plots clearly indicated that it made sense to add number of accents to the independent variable "musician" and the covariate "number of distinct IOIs".

If complexity depends on entropy, and entropy depends on number count, then we might naively expect complexity to depend on number count, too. To resolve this conundrum, I use the variance inflation factor (V.I.F.) introduced in Section 2. Recall that in the model complexity ~ entropy + number/density of notes + order effects + random intercept for subject and excerpt, the V.I.F. tests for multicollinearity in the predictor variables, and finds none. Entropy and number/density of notes (which are approximately the same in the experimental setup used here) tell us different things about the data. We need not concern ourselves with the propagation of the choice between number count and number density from the entropy calculation to the perceived complexity calculation. In other words, we don't need to know what the exact relationship is between number (density or count) and entropy in the perceived complexity calculation. All we need to know is revealed by the estimated marginal means procedure for computational results and the linear mixed effects regression procedure for experimental results.

It should be noted, however, that the absence of triplet- or sixteenth-note-based rhythms or ornamental straight-eighth-notes in the experimental excerpts may impact the application of note density to any theorizing about how note density affects the perception of rhythmic complexity. For example, a subsequent experiment might indicate that quarter-note triplets impact one's perception of rhythmic complexity more, or at least differently, than eighth-notes do.

**3.7 Conclusion**

In this section I tested the hypothesis that computed entropy depends on musician in a corpus of 88 solos by Armstrong, Hawkins, Young, Christian, and Parker using the estimated marginal means technique with number of IOIs and number of accents as covariates. Using this technique obviated the need for normalization or minimum sample size. Furthermore, the added variable plot technique strongly supported the inclusion of these covariates.

Results showed that solos by Young were significantly more entropic than solos by Armstrong or Parker, and *probably* more entropic than solos by Christian (the presence of researcher error made it impossible to ascertain the results regarding

Christian for sure). Furthermore, two hypotheses regarding chronological trends were shown not to be supported by entropy calculations.

The lack of contrast in entropy between Hawkins and Parker was explained as probably being due to the greater dynamic accent density in the solos of Hawkins, and the greater contour accent density in the solos of Parker.

Finally, it was shown that it is not necessarily to know whether listeners respond to note *number* or note *density* in order to understand the results found here.

## 4 Entropy and Melodic Embellishment

A vital part of the jazz tradition is the embellishment of standard songs to become jazz. Examples can be found in the heads corresponding to the improvisations in the corpus used here, such as Young's renditions of "All of Me" and "Tea For Two." In other cases, such as Hawkins's famous 1939 rendition of "Body and Soul" (see chapter 6), or Armstrong's renditions of "Stardust" (1931) and "What a Wonderful World" (1967), the embellishments are so extreme as to render the heads part of the solos themselves.

Melodic embellishment is one of the four types of improvisation enumerated by Henry Martin in his study of thematic improvisation in Charlie Parker's playing (1996, 34), based on classifications devised by Kernfeld and Hodier – namely, paraphrase improvisation[14]. The question of when a paraphrase becomes a freer type of improvisation is difficult to answer; apparently, the dividing line is based purely on intuition. While this answer is somewhat unsatisfactory, it is also inevitable: musicians

[14] The other three types are chorus phrase improvisation (based on the form and harmonic structure of the head only), motivic improvisation (based on motives from the head), and formulaic improvisation (based on formulas used by the soloist throughout his or her oeuvre).

often reference the melody, to greater or lesser degrees, during improvisation. For example, Hawkins begins his solo on "Epistrophy" by quoting the melody verbatim for eight bars, before developing the rhythmic motives of which the head is composed and producing a varied and fascinating solo. In the present context, melodic renditions that substantially alter the melody have been considered to be solos or parts of solos, and have contributed to the evaluation of entropy, while those that do not substantially alter the melody have not been considered part of an excerpt to be evaluated for entropy. Charlie Parker's rhythmically complex heads were not included in calculations of entropy.

Here I examine several examples of melodic embellishment, namely, embellished versions of "Tea For Two," "All of Me," and" What a Wonderful World," by Lester Young, Doris Day, Ella Fitzgerald, Frank Sinatra, and Louis Armstrong.

"Tea For Two," as written, has very little rhythmic variety. In fact, almost three quarters of the song follow either the template shown in Figure 16 (a), or its close rhythmic relative, Figure 16 (b). The rest of the song consists of whole notes. In all cases, the implied accents are purely metrical, and fall on beats one and three. Not surprisingly, the calculated entropy for this song as written is very low: 0.32, to be exact.

I examined three renditions of this melody, two by singers – Doris Day and Ella Fitzgerald – and one by Lester Young. Somewhat surprisingly, at least in the case of Doris Day, for whom we might expect lower rhythmic unpredictability than for Fitzgerald or Young, the calculated entropies for these three renditions of the song were quite similar: 2.71 for Fitzgerald, 2.75 for Day, and 2.90 for Young.

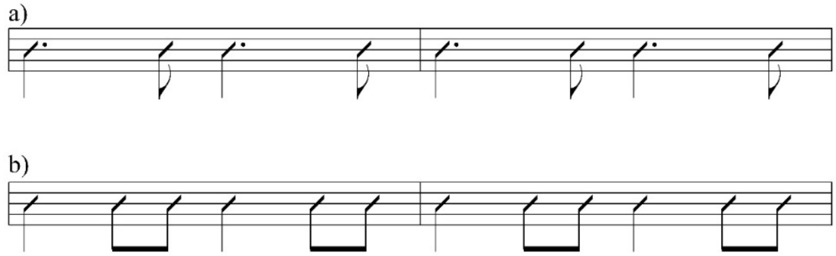


**Figure 16.** Rhythms upon which “Tea For Two” is based

“All of Me” has an even lower entropy: every bar begins with an accented note, and accents do not fall on any beats other than downbeats. Thus, the entropy is, in fact, zero! The entropies corresponding to versions by Frank Sinatra, Ella Fitzgerald, and Lester Young, on the other hand, occupy a narrow band of +/– 1.8% centered on $H$ =3.54: Sinatra has $H = 3.477$, Fitzgerald $H = 3.595$, and Young $H = 3.606$. The renditions by Sinatra and Fitzgerald were taken to be the first of two melody choruses, in order to facilitate comparison with the Young melody chorus; had the second melody choruses been used, the entropies most likely would have been greater.

Note, too, that Sinatra employs a subtle behind-the-beat phrasing, not enough to warrant shifting any rhythms by a sixteenth note, as Louis Armstrong is wont to do, e.g. in “What a Wonderful World” m. 23, where beat one is shifted to the second sixteenth note of beat one, but just enough to be perceptible (Figure 17).

Finally, I compare one of Louis Armstrong’s 1967 recordings of the aforementioned “What a Wonderful World” (Figure 17) to the sheet music version. I find that, for the sheet music, $H = 1.517$, and for the Armstrong recording, $H = 3.877$.

Armstrong
What a Wonderful World


**Figure 17.** Transcription of one of Louis Armstrong's version of "What a Wonderful World"

## 5 Mutual Information and Soloist-Accompanist Interaction

Here I use the concept of mutual information, as defined previously, to analyze the interactions between Charlie Parker and some of the pianists who accompanied him. This is a sensible approach, since we expect the placement of chords by the accompanist to reflect the placement of accents by the soloist and vice-versa; mutual information appears to be the perfect tool to represent these interactions.

There are several possible ways to implement the calculation of mutual information for strings of saxophone accents and their corresponding chord onsets. For simplicity, I restrict myself to the case in which IOI's between saxophone accents and between chord onsets are subdivided solely by eighth notes; that is, I omit solos in which triplet, sixteenth note, or straight-eighth-note subdivisions are used in either the soloist or accompanist parts.

Given this restriction, I subdivide the solo accent and chord onset streams into half-measure – that is, four eighth note – groups. Next, I label each group – in both the soloist and accompanist parts – with a number from 1 to 16, according to the 16 possible combinations of four eighth notes or eighth note rests. (There are two possibilities for each of the four positions in each group, thus 2*2*2*2 = 16). I calculate the joint distribution by counting the number of occurrences of each pair of labels (in the soloist and accompanist parts) and dividing by the number of half-measure groups. I calculate the marginal distributions by simply counting the number of occurrences of each label within first the soloist, then the accompaniment parts, and dividing by the number of half-measure groups. Thus I am able to calculate MI scores using Equation 9.

I begin by considering comping rhythms that result in low MI scores. For an extreme but not far-fetched example, consider comping rhythms with zero entropy. When either distribution in a mutual information calculation has zero entropy, the resulting mutual information is zero as well, since knowledge of one variable has no effect on the knowledge of the other.

Recall that zero entropy rhythms have just one *IOI* value. So, for example, whole notes in the accompaniment part starting on beat one of every measure would yield zero MI, as would quarter notes on every beat. While these rhythms are unlikely to appear in a jazz accompaniment, eighth notes on the "and" of two and the "and" of four can frequently be found in the playing of pianist Red Garland; this, too, would yield zero MI.

Next consider a more complex but still simplistic comping rhythm: the previously discussed "Charleston" rhythm (see Figure 6). Here the MI scores are non-zero, but still low – as compared to the MI scores obtained using the actual comping rhythms (see Table 9), by a factor  ranging from 2.63 to 22.65. Thus, as expected, MI reflects to a certain extent the interaction between soloist and accompanist.

Next I calculate MI scores using randomly generated comping rhythms with the same number of chords as in the actual rhythm, and found that, with one exception, they were higher than the scores calculated using the actual comping rhythms. Factors ranged from 0.98 to 2.0. This makes sense since a randomly generated accompaniment would not contain many repeated segments, so knowledge of one's place in the accompaniment would yield a great deal of information about one's place in the solo.

| Title | MI actual | MI Charleston | MI Random | actual/Charleston | random/actual |
|---|---|---|---|---|---|
| KoKo | 0.07 | 0.03 | 0.15 | 2.63 | 2.00 |
| Bloomdido | 0.25 | 0.05 | 0.28 | 5.49 | 1.12 |
| Dewey Square | 0.33 | 0.12 | 0.53 | 2.72 | 1.61 |
| Donna Lee | 0.34 | 0.04 | 0.34 | 9.08 | 0.98 |
| Yardbird Suite | 0.37 | 0.10 | 0.52 | 3.87 | 1.38 |
| Crazeology | 0.40 | 0.09 | 0.43 | 4.46 | 1.08 |
| Cheryl | 0.44 | 0.06 | 0.50 | 7.49 | 1.14 |
| Ornithology | 0.50 | 0.02 | 0.61 | 27.65 | 1.21 |
| Bongo Beep | 0.59 | 0.19 | 0.73 | 3.06 | 1.23 |
| Au Privave | 0.63 | 0.08 | 0.69 | 7.87 | 1.09 |

**Table 9.** Mutual Information for ten Parker excerpts: actual, calculated using a Charleston comping rhythm, and calculated using a random comping rhythm. Ratios of actual MI to MI calculated using Charleston rhythm, MI calculated using random comping rhythm to actual MI.

Finally, I point out a possible reason for the MI score for “KoKo” (recorded 11/26/1945) being lower than the other MI scores. The original pianist booked for this recording session was Bud Powell, but he had to cancel at the last minute; as a replacement, the little known pianist Argonne Thornton (who later changed his name to Sadik Hakim) was brought in. There is disagreement as to whether it was Thornton who played piano on “KoKo,” or the trumpeter Dizzy Gillespie filling in on his second instrument. It is unlikely that either of these accompanists would have possessed the pianistic agility displayed by any of the other accompanists represented here (Duke Jordan, Al Haig, Dodo Marmarosa, Bud Powell, Thelonious Monk, or Walter Bishop Jr.). Thus the low MI score; note, too, that the ratio of actual MI to Charleston MI is the lowest of the ten solos evaluated. So this is an important result.

Overall, then, mutual information appears to be a useful measure of the interaction between soloist and accompanist.

## 6 Conclusion

Within roughly ten years of its founding as a discipline, information theory had been applied to music by several authors. One reason for this is surely that the formula for Shannon entropy – the central concept of information theory – is simple, and can be used with any random variable described by a known probability distribution. Another, more enlightening, reason is provided by Meyer (1957), who posits that, in music, the phenomenon of expectation, specifically *thwarted* expectation, correlates with both information and musical meaning.

The present work explored the relationship of entropy, and of the related concept of mutual information, to jazz rhythm in several ways, while making a brief foray into the study of rhythmic periodicity and syncopation as they pertain to the perception of rhythmic complexity.

First of all, it demonstrated that entropy derived from IOIs between accented notes in jazz solos is a significant factor in the perception of rhythmic complexity for jazz rhythms constructed solely of multiples of eighth-notes and eighth-note rests. This conclusion was reached by means of an experiment that asked fifteen music majors to rank eighteen short rhythmic excerpts for complexity on a scale from one to seven. Results were analyzed using a mixed-effects multiple regression model with the Satterthwaite approximation. Multiple predictor variables were used, in addition to entropy: number of notes (which was roughly the same as note density because excerpts were approximately the same length), two variables quantifying the periodicity of test rhythms (or the lack thereof), syncopation, order effects (in other words, the effects of listening to the experimental excerpts in different orders), and jazz experience level. In the course of identifying a model that included entropy, number of notes, and excerpt order as significant factors, I also identified a model that included lack of periodicity, syncopation, number of notes, and order effects as significant factors in the perception of rhythmic complexity. It is likely that entropy was mediated by periodicity in its effect on rhythmic complexity ratings.

Next, using a corpus of 88 transcribed solos by Louis Armstrong, Coleman Hawkins, Lester Young, Charlie Christian, and Charlie Parker, I demonstrated that solos by Young are more entropic than those by Armstrong, Parker, and probably Christian, but not Hawkins. I arrived at this conclusion by isolating dynamical accented notes and calculating probability distributions based on the inter-onset intervals between accented notes, calculating the entropy of each excerpt, and using the estimated marginal means technique with the Bonferroni correction to make pairwise comparisons between musicians. Two covariates were used in this calculation: number of distinct IOIs and number of accents. The utility of these choices was confirmed using the added variable plot technique; including these covariates obviated the need for normalizing the computed entropies or insisting on minimum sample sizes.

Two commonly received notions about the solos of Louis Armstrong and Lester Young were rebuked using the estimated marginal means technique as applied to entropy. The lack of difference in entropy between Parker and Hawkins was explained by the greater dynamic accent density in the solos of Hawkins and the greater contour accent density in Parker. Finally, it was shown that it is not necessary to know whether listeners respond to note *number* or note *density* in order to understand the results obtained here.

The phenomenon of melodic embellishment, central to jazz, was explored using entropy. Embellished versions of standard songs were uniformly found to have higher entropies than their sheet-music counterparts.

Mutual information was used to study the interaction between Parker and his piano accompanists in a limited corpus of ten transcribed solos. Mutual information between Parker and his accompanists was found to be greater than that between Parker and a repeating Charleston rhythm accompaniment, and less than that between Parker and a random accompaniment. The latter result makes sense because knowledge of one's place in a random accompaniment yields a high amount of information about one's place in the solo. The lowest mutual information value and lowest ratio of mutual information between Parker and his accompanist to the mutual information between Parker and a Charleston accompaniment occurs on "KoKo"; this may be due to the fact that the accompanist on the date, either Argonne Thornton (Sadik Hakim) or Dizzy Gillespie, was probably less pianistically agile than Parker's usual accompanists.

Thus, this research has indicated that entropy is probably a significant factor in the perception of rhythmic complexity, that lack of periodicity probably mediates between entropy and perceived complexity, that entropy depends on musician in a corpus of 88 solos by five great jazz musicians, that entropy is a useful tool for understanding melodic embellishment, and that mutual information is a useful tool for studying the interaction between soloist and accompanist.

It points the way toward at least four avenues for future exploration.

First and foremost, this research calls for experimental studies of entropy as a significant factor in the perception of rhythmic complexity for a more general class of rhythms, those including, triplets, sixteenth notes, and ornamental straight eighths. It is possible that the conclusions reached in the present study might be changed in a more general experimental context. For example, the presence or absence of triplets, sixteenth notes and ornamental straight eighths might overwhelm the effects of entropy on perceived rhythmic complexity. Or the effects of number or density of notes on complexity might interact with the types of rhythm included: perhaps the number of quarter-note triplets might have a stronger effect on perceived rhythmic complexity than the number of eighth-notes alone.

This research calls for studying carryover effects in such a way as to isolate the effects of predictor variables on subsequent complexity ratings. It seems intuitive that an excerpt with high entropy followed by an excerpt with low entropy might artificially deflate the subject's reaction to the second excerpt. It is unlikely, however, that this would significantly change the experimental results obtained here, if the negligible magnitude of the carryover effects isolated here is any indication.

It calls for computational research with more excerpts and more musicians. An equal number of excerpts should be transcribed for all musicians, and this should be as high as possible. And adding other musicians to the solos might help identify what exactly the factors are that lead to low or high entropy as it was here defined. Would Lester Young's solos be more or less entropic than John Coltrane's? Or Benny Goodman's? An alternative to transcribing more solos would be to explore extant transcriptions of jazz solos available online. This might be an efficient way to expand the corpus.

And finally, it calls for corpora that are larger and include musicians other than only Charlie Parker in the study of mutual information. The evidence is, however, that mutual information is a useful tool in studying soloist/accompanist interaction. Perhaps there are specific issues that can be explored using mutual information, similar to the anecdote regarding Sadik Hakim/Dizzy Gillespie.

Thus, building on previous work in the field of information theory as applied to classical and jazz music, and using both experimental and computational techniques, this work adds a valuable new perspective on the study of rhythmic complexity in jazz.

## Appendix A

## Syncopation Metric of Longuet-Higgins and Lee

Longuet-Higgins and Lee (1984) present a simple but elegant metric (the "LHL" metric) for describing syncopation. Since this is applied in the experimental portion of this study, I will explain it here in some detail.

The LHL metric, when applied to eighth-note rhythms in 4 4 time, works by assigning each part of the measure a number: 0 for the downbeat, –1 to beat three, –2 to the off–beat quarter notes (two and four), and –3 to the four off-beat eighth notes (the "ands" of one, two, three, and four). A syncopation is said to occur when a note sounds before a tied note or rest having a greater (more positive) value than the sounding note. The "weight" of the syncopation is equal to the greater value minus the lesser value, which will necessarily be a positive number. Note that a single sounded note can correspond to multiple syncopations, according to the number of rests and/or tied notes following the sounded note.

Several examples are shown in Figure 18: (a) A note that sounds on the "and of two" (i.e., its note attack begins an eighth–note after the second beat) and is tied to a quarter note on beat three will count as a syncopation, since the value assigned to the tied note (–1) is greater than the value assigned to the sounding note (–3). The weight of the syncopation will be –1 – (–3) = 2. If there is a rest on beat four, there will be an additional syncopation of weight 1, obtained from the difference between the value assigned to beat four (–2) and the value assigned to the sounding note (–3). (b) A quarter note that sounds on beat four and is tied over to the beginning of the next measure will count as a syncopation, since the value assigned to the tied note (0) is greater than the value assigned to the sounding note (–2). The weight of the syncopation will be 0 – (–2) = 2. (c) A note that sounds on the and of four and is tied to the beginning of the next measure will count as a syncopation with weight 0 – (–3) = 3. If there is a rest on beat two, there will be another syncopation, this one of weight 1, obtained from the difference between the value assigned to beat two (–2) and the value assigned to the sounding note on the and of four (–3).

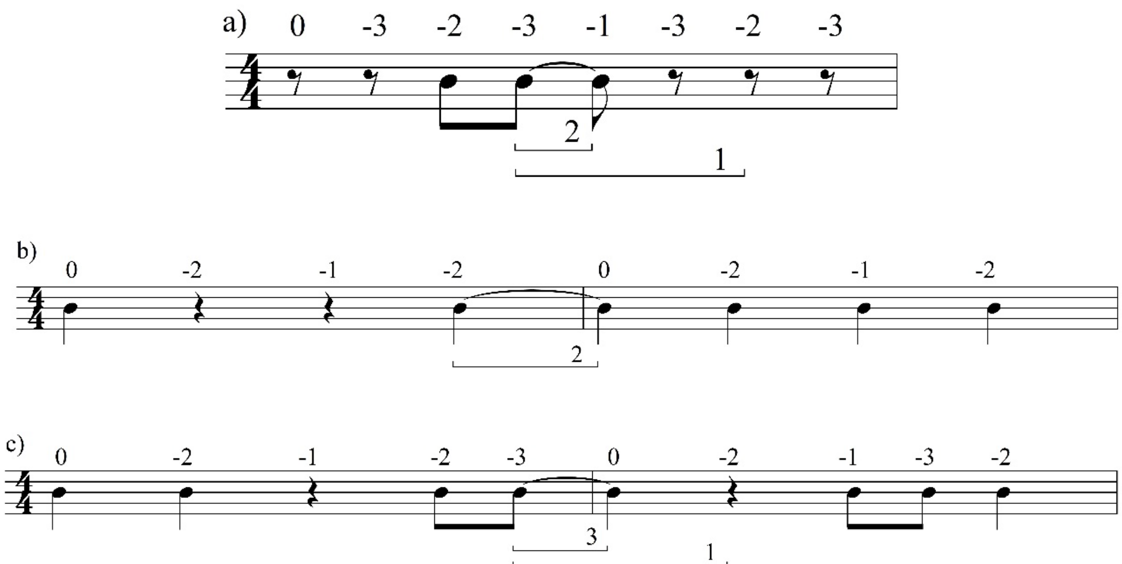


**Figure 18.** Three examples of the LHL metric: a) a total of three (two syncopations initiated by the same note); b) a total of two (one syncopation only); a total of four (two syncopations initiated by the same note)

For purposes of this study, LHL scores will be divided number of note onsets to obtain the “LHL quotient” or “*LHLQ*.”

## Appendix B

## Catalog of Excerpts

| Name | Date | Title | TS | IOIs | acc's | Entropy |
|---|---|---|---|---|---|---|
| Armstrong | 26 | Cornet Chop Suey pt. 2 | | 14 | 84 | 3.138 |
| Armstrong | 27 | Potato Head Blues pt. 1 | | 12 | 31 | 3.24 |
| Armstrong | 27 | Potato Head Blues pt. 2 | | 12 | 100 | 3.03 |
| Armstrong | 29 | Mahogany Stomp pt. 1 | | 8 | 19 | 2.795 |
| Armstrong | 29 | Mahogany Stomp pt. 2 | | 8 | 14 | 2.815 |
| Armstrong | 29 | Mahogany Stomp pt. 3 | | 10 | 54 | 2.58 |
| Armstrong | 31 | Stardust pt. 1 | ✓ | 15 | 39 | 3.491 |
| Armstrong | 31 | Stardust pt. 2 | ✓ | 18 | 45 | 3.327 |
| Armstrong | 31 | Stardust pt. 3 | ✓ | 23 | 55 | 4.174 |
| Armstrong | 38 | I Double Dare You | | 15 | 65 | 3.11 |
| Armstrong | 43 | I Never Knew | ✓ | 12-13 | 33-35 | 3.201 |
| Armstrong | 44 | I'm Confessin' (That I Love You) | | 14 | 43 | 3.208 |
| Armstrong | 47 | It Takes Time | ✓ | 9 | 36 | 2.785 |
| Armstrong | 53 | The Gypsy | ✓ | 19 | 71 | 3.305 |
| Armstrong | 56 | A Foggy Day | ✓ | 11 | 32 | 3.02 |
| Armstrong | 63 | Hello Dolly | | 13 | 39 | 2.803 |
| Armstrong | 67 | What a Wonderful World | ✓ | 23 | 79 | 3.877 |
| Hawkins | 26 | The Stampede | | 11 | 72 | 2.9 |
| Hawkins | 33 | Talk of the Town | | 16 | 76 | 3.393 |
| Hawkins | 39 | Body and Soul | ✓ | 20 | 236 | 2.773 |
| Hawkins | 40 | Bouncin' With Bean | | 12 | 86 | 3.027 |
| Hawkins | 40 | My Blue Heaven | | 9 | 75 | 2.046 |
| Hawkins | 41 | Disorder at the Border | | 12 | 108 | 2.724 |
| Hawkins | 44 | Flyin' Hawk | ✓ | 15 | 100 | 3.336 |
| Hawkins | 50 | Ballade pt. 1 | ✓ | 15 | 57 | 3.238 |
| Hawkins | 50 | Ballade pt. 2 | ✓ | 15 | 55 | 3.435 |
| Hawkins | 54 | Lullaby of Birdland | ✓ | 16 | 55 | 3.341 |
| Hawkins | 57 | Blues for Tomorrow | ✓ | 17 | 215 | 2.002 |
| Hawkins | 57 | Epistrophy | | 18 | 82 | 3.665 |
| Hawkins | 62 | Satin Doll | ✓ | 24 | 215 | 2.862 |
| Hawkins | 62 | Wanderlust | ✓ | 21 | 62 | 3.794 |
| Hawkins | 63 | Just Friends | | 15 | 79 | 3.122 |
| Young | 36 | Lady Be Good | | 16 | 99 | 3.465 |
| Young | 36 | Shoe Shine Boy | | 20 | 86 | 3.656 |
| Young | 38 | Back in Your Own Backyard | | 13 | 51 | 3.271 |
| Young | 38 | When You're Smiling | | 12 | 44 | 3.071 |
| Young | 42 | Body and Soul pt. 1 | ✓ | 19 | 73 | 3.616 |
| Young | 42 | Body and Soul pt. 2 | ✓ | 14 | 52 | 3.392 |
| Young | 42 | Indiana pt. 1 | ✓ | 18 | 92 | 3.798 |
| Young | 42 | Indiana pt. 2 | | 21 | 92 | 3.557 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Young | 42 | Tea For Two 1942 | | 17 | 113 | 3.189 |
| Young | 45 | D.B. Blues | ✓ | 17 | 87 | 3.64 |
| Young | 46 | It's Only a Paper Moon | | 21 | 163 | 3.328 |
| Young | 47 | Sheik of Araby pt. 1 | | 17 | 77 | 3.624 |
| Young | 47 | Sheik of Araby pt. 2 | | 14 | 46 | 3.463 |
| Young | 47 | Tea For Two 1947 | | 17 | 82 | 3.435 |
| Young | 50 | Blues For Greasy JATP | | 21 | 38 | 4.186 |
| Young | 52 | Ad Lib Blues pt. 1 | | 22 | 132 | 3.441 |
| Young | 52 | Ad Lib Blues pt. 2 | | 19 | 194 | 3.158 |
| Young | 56 | All of Me pt. 2 | | 21 | 117 | 3.692 |
| Young | 56 | Taking a Chance on Love | ✓ | 22 | 134 | 3.822 |
| Young | 56 | You Can Depend on Me | ✓ | 22 | 158 | 3.658 |
| Christian | 39 | Christian Honeysuckle Rose | | 13 | 64 | 2.83 |
| Christian | 39 | Flying Home | | 10 | 61 | 2.789 |
| Christian | 39 | Good Morning Blues | | 17 | 46 | 3.497 |
| Christian | 39 | Rose Room | ✓ | 17 | 60 | 3.043 |
| Christian | 39 | Seven Come Eleven | ✓ | 17 | 43 | 3.638 |
| Christian | 40 | Benny's Bugle | | 13 | 35 | 3.241 |
| Christian | 40 | Gone With "What" Wind | | 13 | 39 | 3.22 |
| Christian | 40 | Grand Slam | | 16 | 36 | 3.557 |
| Christian | 40 | I Can't Give You Anything But… | ✓ | 13 | 34 | 3.002 |
| Christian | 40 | Six Appeal | ✓ | 9 | 36 | 2.695 |
| Christian | 40 | Till Tom Special | | 11 | 20 | 3.003 |
| Christian | 40 | Wholly Cats | | 11 | 33 | 3.097 |
| Christian | 41 | Breakfast Feud | | 13 | 38 | 3.199 |
| Christian | 41 | I've Found a New Baby | | 10 | 46 | 2.828 |
| Parker | 40 | Moten Swing | | 13 | 58 | 2.522 |
| Parker | 40 | Parker Honeysuckle Rose | | 10 | 62 | 2.199 |
| Parker | 40 | Parker Lady Be Good | | 12 | 50 | 2.609 |
| Parker | 41 | Swingmatism | | 11 | 18 | 3.264 |
| Parker | 42 | Hootie's Blues | ✓ | 12 | 32 | 3.011 |
| Parker | 45 | Billie's Bounce | | 17 | 66 | 3.351 |
| Parker | 46 | Yardbird Suite | | 14 | 48 | 3.556 |
| Parker | 47 | Bongo Beep | ✓ | 13 | 40 | 3.216 |
| Parker | 47 | Cheryl | | 18 | 46 | 3.538 |
| Parker | 47 | Crazeology | | 12 | 53 | 2.699 |
| Parker | 47 | Ornithology | | 17 | 43 | 3.461 |
| Parker | 48 | Segment | | 17 | 93 | 3.414 |
| Parker | 49 | Scrapple From the Apple | | 17 | 42 | 3.747 |
| Parker | 50 | Bloomdido | | 17 | 73 | 2.972 |
| Parker | 50 | Mohawk | | 20 | 78 | 3.656 |
| Parker | 51 | Au Privave | | 19 | 44 | 3.952 |
| Parker | 51 | Blues For Alice | | 21 | 54 | 4.033 |
| Parker | 51 | She Rote | ✓ | 22 | 93 | 3.554 |
| Parker | 53 | Chi Chi | ✓ | 21 | 102 | 3.794 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Parker | | 53 | Now’s The Time | ✓ | 25 | 78 | 4.157 |
| Parker | ?? | | Dewey Square | | 16 | 46 | 3.297 |
| Parker | ?? | | Donna Lee | | 21 | 86 | 3.756 |

*TS checked indicates a 2:1 swing ratio and dependence of entropy on swing ratio*

*Multiple excerpts from a single recording reflect intervening passages played by other soloists or intervening ensemble passages*

*For Armstrong/I Never Knew, ranges of num_IOI and num_accent, and average entropy, are given for different dispositions of Armstrong’s “clams”*

*For a detailed list of recording data, please contact the author at dougabrams.jazz@gmail.com*

## Appendix C

## Treatment of Estimated Errors

There were 6,299 accented notes included in the corpus, and thus easily 20,000–30,000 notes altogether. In a manually transcribed corpus of this size, there are bound to be errors. Overall, the error rate – estimated using a random sample of 37 excerpts – was fairly low: insertion, deletion, or translation of accented notes as a percentage of number of accented notes was about 0.5%, and as a percentage of all notes, was lower. There is a potential problem, however, in that the errors are distributed evenly across excerpts; thus, the number of excerpts with minor errors as a percentage of the total number of excerpts – estimated using the same random sample – was much higher: about 40%. Given that files discovered with errors were corrected, this means that in the corpus as a whole, there were probably about 20 files with minor errors after review and correction.

Once errors were discovered, I used four methods of quantifying them. Percentage errors in entropy, number of IOIs, and number of accents were identified, and pairwise *p* values were calculated before and after correction. For a limited corpus of twelve error–containing excerpts, Table 10 gives the mean, median, and standard deviation of percentage error in entropy, percentage error in number of IOIs, and percentage error in number of accents. Note that the number of IOIs exhibits the largest average percent error, since making a small change to the data can result in a change in number of IOIs of one or even two, both of which are relatively large in terms of number of IOIs. Finally, including the twelve error-containing excerpts mentioned above and correcting them one by one did not qualitatively change the pairwise *p* values.

While *p* values are relatively insensitive to simple errors, a review of the data indicates that adding or deleting excerpts for various reasons has a more pronounced effect. In the course of analyzing the data, I realized that some excerpts had to be deleted for reasons having to do with the properties of the swing eighth-notes used, either because they were too close to straight-eighths, or because it was too difficult to ascertain whether or not the eighth-note pairs had a 2:1 ratio. For example, for a data set with five error-containing excerpts and with just one more excerpt than the final data set, the Christian/Young *p* value was 0.0406, while for the same data set with *four* more excerpts than the final data set, the *p* value was 0.0603. Furthermore, a data set from early in the experiment with many more error-containing excerpts (27) *and* with four more excerpts than the final data set yielded a *p* value for the Christian/Young pairwise comparison of 0.0541. For all data sets examined, the only *p* values that differ qualitatively from those obtained from the final set are those corresponding to the Christian/Young comparison, and these were at least "marginally significant" (0.0518–0.0627). In any case, it appears as though adding or deleting excerpts has a greater effect on *p* values than the presence or absence of simple errors.

Having settled on a final corpus, and estimating that no more than twenty excerpts contain errors, I argue that the *p* values for Christian/Young and for all other pairwise comparisons are valid. Even if there are twenty error-containing excerpts, it appears that correcting errors drives the Christian/Young *p* value down, therefore not affecting significance. There is, however, a chance that the presence of unknown errors may have an unpredictable effect on the Christian/Young *p* value, perhaps resulting in marginal significance or even insignificance. In this case, the Armstrong/Young and Parker/Young $p$ values still indicate significant pairwise comparisons.

| Quantity/Statistic | Mean | Median | Standard Deviation |
|---|---|---|---|
| Entropy | 1.603 | 1.1 | 1.213 |
| Number of IOIs | 3.683 | 4.7 | 3.510 |
| Number of accents | 0.75 | 0.95 | 0.722 |

**Table 10.** Statistical descriptions of three measures of percentage error: Entropy, Number of IOIs, and Number of accents for a corpus of 12 error-containing excerpts

## REFERENCES


Bell, Chip. "Algorithmic music composition using dynamic Markov chains and genetic algorithms."*Journal of Computing Sciences in Colleges* 27, no. 2 (2011): 99–107.

Christensen, Thomas, ed. *The Cambridge history of Western music theory*. Cambridge University Press, 2010

Cohen, Joel E. "Information theory and music." *Behavioral Science* 7, no. 2 (1962): 137–163.

Cover, Thomas M., and Joy A. Thomas. "Elements of information theory 2nd edition (Wiley series in telecommunications and signal processing)." (2006).

De Fleurian, Remi, Oded Ben-Tal, Daniel Müllensiefen, and Tim Blackwell. "Comparing perceptual and computational complexity for short rhythmic patterns." *Procedia-Social and Behavioral Sciences* 126 (2014): 111-112.

DeFleurian, Remi, Tim Blackwell, Oded Ben-Tal, and Daniel Müllensiefen. "Information-theoretic measures predict the human judgment of rhythm complexity." *Cognitive Science* 41, no. 3 (2017): 800-813.

DeVeaux, Scott. *The birth of bebop*. University of California Press, 1997.

Devore, Jay L., and Kenneth N. Berk. *Modern mathematical statistics with applications*. New York: Springer, 2012.

Dowling, W. Jay. "Scale and contour: Two components of a theory of memory for melodies." *Psychological review* 85, no. 4 (1978): 341.

Fitch, W. Tecumseh, and Andrew J. Rosenfeld. "Perception and production of syncopated rhythms." *Music Perception* 25, no. 1 (2007): 43–58.

Frieler, Klaus. "Constructing Jazz Lines Taxonomy, Vocabulary, Grammar." *Jazzforschung heute: Themen, Methoden, Perspektiven* (2019): 103–132.

Gillick, Jon, Kevin Tang, and Robert M. Keller. "Machine learning of jazz grammars." *Computer Music Journal* 34, no. 3 (2010): 56–66.

Hiller, Lejaren, and Ramon Fuller. "Structure and information in Webern's Symphonie, Op. 21."*Journal of Music Theory* 11, no. 1 (1967): 60–115.

Huron, David. *Sweet anticipation: Music and the psychology of expectation*. MIT press, 2008.

Joel, Lester. *The Rhythms of Tonal Music*. Carbondale: University of Illinois Press. 1986

Kernfeld, Barry. ed. 1996. *New Grove Dictionary of Jazz.* New York: St. Martin's Press

Knopoff, Leon, and William Hutchinson. "Entropy as a measure of style: The influence of sample length." *Journal of Music Theory* 27, no. 1 (1983): 75–97.

Knopoff, Leon, and William Hutchinson. "Information theory for musical continua." *Journal of Music Theory* 25, no. 1 (1981): 17–44.

Linskens, Erlijn J., and Gijs Schoenmakers. "Music Improvisation Using Markov Chains." *Maastricht University* (2014).

Longuet–Higgins, H. Christopher, and Christopher S. Lee. "The rhythmic interpretation of monophonic music." *Music Perception* 1, no. 4 (1984): 424–441.

Margulis, Elizabeth Hellmuth, and Andrew P. Beatty. "Musical style, psychoaesthetics, and prospects for entropy as an analytic tool." *Computer Music Journal* 32, no. 4 (2008): 64–78.

Meyer, Leonard B. "Meaning in music and information theory." *The Journal of Aesthetics and Art Criticism* 15, no. 4 (1957): 412–424.

Norgaard, Martin, Jonathan Spencer, and Mariana Montiel. "Testing cognitive theories by creating a pattern–based probabilistic algorithm for melody and rhythm in jazz improvisation." *Psychomusicology: Music, Mind, and Brain* 23, no. 4 (2013): 243.

Parker, Charlie. *Charlie Parker Omnibook: for C instruments (treble clef)*. Alfred Music Publishing, 1978.

Patel, Aniruddh D., and Joseph R. Daniele. "An empirical comparison of rhythm in language and music." *Cognition* 87, no. 1 (2003): B35–B45.

Pearce, Marcus, and Geraint Wiggins. "Improved methods for statistical modelling of monophonic music." *Journal of New Music Research* 33, no. 4 (2004): 367–385.

Pinkerton, Richard C. "Information theory and melody." *Scientific American* 194, no. 2 (1956): 77–87.

Quick, Donya, and Kelland Thomas. "A functional model of jazz improvisation." In *Proceedings of the 7th ACM SIGPLAN International Workshop on Functional Art, Music, Modeling, and Design*, pp. 11–21. 2019.

Quinn, Ian. "The combinatorial model of pitch contour." *Music Perception* 16, no. 4 (1999): 439-456.

Roeder, John. "A calculus of accent." *Journal of Music Theory* (1995): 1-46.
Rouse, Tevin. "Stochastic Improvisation of Jazz Solos." (2017).

Sadakata, Makiko, Peter Desain, and Henkjan Honing. "The Bayesian way to relate rhythm perception and production." *Music Perception* 23, no. 3 (2006): 269–288.

Shannon, Claude E. "A mathematical theory of communication." *The Bell system technical journal* 27, no. 3 (1948): 379–423.

Shmulevich, Ilya, and Dirk–Jan Povel. "Complexity measures of musical rhythms."*Rhythm perception and production* (2000): 239–244.

Smith, Leigh M., and Henkjan Honing. "Evaluating and extending computational models of rhythmic syncopation in music." In *ICMC*. 2006.

Snyder, John L. "Entropy as a measure of musical: The influence of a priori assumptions." *Music Theory Spectrum* 12, no. 1 (1990): 121–160.

Temperley, David. *Music and probability*. MIT Press, 2007.

Thul, Eric, and Godfried T. Toussaint. "Rhythm Complexity Measures: A Comparison of Mathematical Models of Human Perception and Performance." In *ISMIR*, pp. 663-668. 2008.

University of Virginia Website: library.virginia.edu/data/articles/introduction-to-mediation-analysis accessed 2/27.24

Victor, Joseph. "Jazz Melody Generation and Recognition." (2012).

Weisberg, Sanford. *Applied linear regression*. Vol. 528. John Wiley & Sons, 2005.

Williams, Martin "Williams, Martin. *The Smithsonian collection of classic jazz*. Smithsonian Collection of Recordings, 1997.

Youngblood, Joseph E. "Style as information." *Journal of Music Theory* 2, no. 1 (1958): 24–35.

Yun, Kristy. "Constraints and Markov chains for a Cognitive Model of Jazz Improvisation: Adding Chords." (2016).