How Jazz Musicians Improvise: The Central Role of Auditory and Motor Patterns

Martin Norgaard
Georgia State University

It is well known that jazz improvisations include repeated rhythmic and melodic patterns. What is less understood is how those patterns come to be. One theory posits that entire motor patterns are stored in procedural memory and inserted into an ongoing improvisation. An alternative view is that improvisers use procedures based on the rules of tonal jazz to create an improvised output. This output may contain patterns but these patterns are accidental and not stored in procedural memory for later use. The current study used a novel computer-based technique to analyze a large corpus of 48 improvised solos by the jazz great Charlie Parker. To be able to compare melodic patterns independent of absolute pitch, all pitches were converted to directional intervals listed in half steps. Results showed that 82.6% of the notes played begin a 4-interval pattern and 57.6% begin interval and rhythm patterns. The mean number of times the 4-interval pattern on each note position is repeated in the solos analyzed was 26.3 and patterns up to 49-intervals in length were identified. The sheer ubiquity of patterns and the pairing of pitch and rhythm patterns support the theory that preformed structures are inserted during improvisation. The patterns may be encoded both during deliberate practice and through an incidental learning processes. These results align well with related processes in both language acquisition and motor learning.

Received: October 22, 2012, accepted May 15, 2013.

Key words: improvisation, corpus research, pattern, Charlie Parker, computer-assisted musical analysis

It has recently been suggested that the study of creativity needs to be divided into various areas according to the cognitive processes involved (Dietrich & Kanso, 2010). One such area is the study of novel output created in real time under certain constraints. Examples of such behavior would be the novel movements of an expert basketball player bypassing defenders from the opposing team, the actual wording of a presenter using an outline, or the notes of a musical improvisation. All these behaviors are created in real time in which revision is not an option. Should errors occur they must be incorporated into upcoming output as they cannot be undone. All these behaviors are guided by rules. The basketball player must execute moves that are physically possible and that are allowed within the rules of the game, the presenter will follow the outline using words and phrases guided by grammatical rules, and the improviser will use notes that fit within the given musical context.

It is well known that only expert practitioners are indeed able to create this output seamlessly. A beginner basketball player will drop the ball just like the novice presenter will stumble over words and lose his train of thought. Through training, execution becomes more fluid and errors subside. A large part of this learning process is mastering the motor movements themselves (e.g., Schmidt & Lee, 2011) but the behaviors described above add another level of complexity. How are discrete movements combined to create a novel output that uniquely fits a given situation? In the case of musical improvisation the problem can be stated as follows: Performance of both composed and improvised music involves learned movements. However, during musical improvisation the exact configuration of those movements is determined in the moment. How is this accomplished? What information is stored in the brain that enables this complex behavior? One theory posits that memorized motor patterns form the basis for the improvised output (Pressing, 1988). A competing theory emphasizes learned rule-based procedures (Johnson-Laird, 2002). The current study explores this question by analyzing a large corpus of improvisations by jazz great Charlie Parker.

It is well documented that jazz improvisations contain repeated melodic figures often referred to as patterns (Berliner, 1994; Finkelman, 1997; Owens, 1974). One explanation is that the artist simply inserts pre-learned structures (Pressing, 1988). Pressing divided improvisations into collections of note groupings that he labeled "events." Each event is triggered by a creative intention in the form of a mental schema that contains a cognitive

Music Perception, volume 31, issue 3, pp. 271–287, issn 0730-7829, electronic issn 1533-8312. DOI: 10.1525/mp.2014.31.3.271

image of sound and corresponding motor realization. As the mental schema is executed, the improviser compares the intention with the actual performed output through various feedback loops. Subsequent events often share features with preceding events, resulting in a related set of note groupings. The improviser may also choose to interrupt the flow by initiating an event that is completely unrelated to the preceding event.

An alternate explanation is that the patterns simply appear by chance as the improviser uses learned procedures that follow the rules of tonal jazz (Johnson-Laird, 2002). Johnson-Laird argues two main constraints are enough to determine the pitches of the improvised output in tonal jazz. The first is the underlying chord progression and procedures related to the use of scales, chord tones, and passing tones. The second is contour considerations. The selected pitches are inserted into prototypical jazz rhythms. Crucially, the entire process does not require storage of intermediate results explaining how subconscious procedures may control the improvisational output without the use of working memory. According to this explanation "instead of a list of fragments of rhythms, motifs, and so on, the algorithms described make use of rules." (Johnson-Laird, 2002, p. 440). Acknowledging that improvised solos often contain motifs that are developed later in the solo, Johnson-Laird argues that extemporized salient phrases may be stored in long-term memory and then used later in the same solo. However, according to Johnson-Laird, the improviser does not develop a memorized library of melodic figures that is accessed during all improvisations.

Norgaard (2011) suggests that both processes occur at different times. This qualitative investigation was based on participants' comments concerning their thinking during an improvisation. Immediately following an improvised performance, participants were interviewed as they were listening to a recording and looking at notation of the solo. The two most often cited strategies for generating note content described by the participants were inserting well-learned ideas from memory and choosing notes based on a harmonic priority. Following the first strategy, participants described how material was retrieved from memory and concatenated "like having a bunch of Legos and how the Legos can fit together" (Norgaard, 2011, p. 119). At times they described inserting the exact pattern stored in memory but most times they acknowledged that the particular pattern was changed to fit the current context. The second strategy was described as "weaving through the changes" (Norgaard, 2011, p. 119) by connecting chord tones placed on strong beats with various scalar and chromatic passing tones. Though the improvisers' descriptions of their own thinking may be different from the actual cognitive processes used, it is extraordinary how well these two strategies align with the two competing theories mentioned above.

One recent study investigated whether learned patterns or procedures guided the improvisational choices of jazz musicians (Goldman, 2012, 2013). Advanced jazz pianists were asked to improvise on a common jazz chord progression (rhythm changes) with either the right or left hand and in both a familiar key (B-flat) and an unfamiliar key (B). Dependent variables included the use of non-diatonic notes and entropy as a measure of melodic predictability. Participants used more diatonic notes and less variation in the unfamiliar key. Goldman interpreted these findings in relation to the generative strategies employed. Forced to improvise in an unfamiliar key, "improvisers used their explicit knowledge of chord tones and scales to improvise" (Goldman, 2012, p. 367) as stored motor programs were not available. Indeed, the improvisers described accessing auditory images of appropriate patterns that they were not able to execute in an unfamiliar key. Therefore, it appears that familiarity with keys may influence whether patterns or procedures are the dominant improvisational strategy.

To more closely examine how improvisational material is created, I first review relevant research on motor behavior and how movements can become integrated with an auditory image. I then describe research that addresses the function of patterns in language learning. Finally, the current research paradigm is outlined and linked to previous analysis of improvisational corpora.

RELATED MOTOR RESEARCH

Pressing's (1988) theory fits well with the idea of a generalized motor program (GMP) as a central memory structure in motor learning, as suggested by the classic schema theory (Schmidt, 1975; for a review, see Shea & Wulf, 2005). The GMP governs the structure of the movement including the sequencing of submovements, relative timing, and relative force. However, the information is stored in an abstract format that is independent of effector use and specifications for absolute timings and force. During execution, a recall schema that includes the specific parameters including effectors needed, absolute timing, and absolute force are scaled to fit the specific situation. Applied to musical improvisation and using Pressing's vocabulary, each event would initiate a GMP that would control pitch and rhythm choices until the following event is initiated. The same GMP used to play a particular pitch and rhythm pattern could be adapted to various tempi (relative timing constant) and could include the same accents whether overall volume was loud or soft (relative force constant).

Building on the original theory, Park and Shea (2005) recently showed that longer movements may consist of individual linked GMPs. "Most interesting, those movement sequences appear to be composed of a number of linked subsequences that form a stable structure for the movement and require more than 3 s to complete" (Shea & Wulf, 2005, p. 98). Using a lever moving task, a sequence divided into 10 or 16 elements was investigated after one or four days of practice. After one day of practice the 10-element sequence was organized into fewer subsequences and stored in a relatively abstract format as shown by an effector transfer test. After four days of practice the entire 16-element sequence was performed without obvious transitions but this information appeared effector dependent. The experiment showed that practice is used to chunk submovements into larger sequences. It appears the longer the sequence, the more specific movement information is stored. This aligns with the notion that initial practice involves the effector independent GMP but later practice stores the actual recall schema which is less malleable (Shea & Wulf, 2005).

AUDIO-MOTOR INTEGRATION

As musicians learn the motor movements necessary to play an instrument, they develop links between those movements and the sounds they produce (Baumann et al., 2007; Drost, Rieger, Gunter, & Prinz, 2005). Subsequently, in a button pressing task in which the button placement was either related or unrelated to the pitch height of simultaneously presented tones, related movement and pitch mappings facilitated faster movement initiation (Keller & Koch, 2008). Furthermore, movement initiation time was positively correlated with participants' musical experience. The authors speculate that the musicians in the study pre-planned the entire movement prior to movement initiation by imagining the related pitch sequence. The ability to imagine the entire sequence "may reflect a stronger tendency in musicians than in nonmusicians to represent sequential action–effect tones as melodic chunks rather than isolated events." (Keller & Koch, 2008, p. 290).

One explanation for this tight coupling of auditory and motor information is that both types of information engage the same brain regions (Baumann et al., 2007; Chen, Rae, & Watkins, 2012; Lahav, Saltzman, & Schlaug, 2007). Indeed, nonmusicians trained to play a melody on the piano show increased activation in a motor related network during subsequent brain imaging while listening to the just learned sequence as compared to listening to a similar but unpracticed melody (Lahav et al., 2007). This shows that audio-motor integration for particular movements and pitches can be developed over a short training period lasting only five days.

It is possible that expert musicians not only bind actions with their contingent musical sounds, but also develop a kind of motor grammar that facilitates their ability to connect actions in musical sequences (Novembre & Keller, 2011). In this study, pianists attempted to imitate a chord progression by watching silent videos of another pianist's hands. Participants made more imitation errors and had slower reactions when the final chord in the progression was harmonically incongruent with the one preceding it, presumably because this movement sequence was less familiar. According to the authors, developing a movement grammar that is linked to syntactical musical rules could "become particularly important for composition and improvisation" (Novembre & Keller, 2011, p. 1242).

LEARNING PATTERNS

Research findings in the area of language learning suggest that the developing improviser might learn a large number of patterns by listening and segmenting existing musical material. Indeed, the ability of infants to segment perceived speech into patterns is central to language learning. Furthermore, it appears this patterning mechanism is not domain-specific and could therefore shape how music is perceived and produced. An intriguing line of research shows that statistical learning processes are used by infants to segment speech streams (Lew-Williams & Saffran, 2012; Saffran, Aslin, & Newport, 1996; Saffran, Newport, Aslin, Tunick, & Barrueco, 1997; Saffran, 2003; Saffran, Pollak, Seibel, & Shkolnik, 2007). This "statistical learning refers to the domain-general ability to extract structure from patterned input" (Lew-Williams & Saffran, 2012, p. 241). Infants use both forward and backwards probabilities to identify word boundaries (Hay, Pelucchi, Estes, & Saffran, 2011). This way children learn that "pret-ty" and "ba-by" are words as the probability that their syllables appear together is higher than the probability that "ba" follows "ty" (Saffran, 2003). As infants segment speech into possible syllable patterns, these patterns then become easier to associate with objects in the environment (Hay et al., 2011). It follows that these speech sound patterns must be stored in a memorized library.

The same statistical learning mechanism has been shown to be active during a pitch patterning task (Saffran, Johnson, Aslin, & Newport, 1999). In this experiment, both adults and infants were exposed to a continuous auditory stream in which embedded pitch patterns were hidden between random pitches. In order to learn the patterns, participants had to use statistical cues based on the fact that some pitches co-occurred more often. Both infants and adults were able to discern these patterns, suggesting that statistical learning is active during music perception. Furthermore, this learning process appears incidental as participants do not have to focus on the input to discover the patterns (Saffran et al., 1997). This domain general pattern learning mechanism could help explain how improvisers following Pressing's theory develop a library of patterns.

PREVIOUS RESEARCH ON PATTERNS IN JAZZ

To investigate the use of patterns and procedures during musical improvisation, the current study analyzed a large corpus of improvised material by the jazz great Charlie Parker. The statistical analysis of a large corpus of compositions has been used to gather information about the compositional process (Temperley, 2007). In one study, patterns in the harmonic structure of rock songs were analyzed using a corpus of 100 songs to illuminate compositional principles (de Clercq & Temperley, 2011). It appears the current investigation is the first time a large corpus of improvised solos has been analyzed using computer algorithms to investigate the underlying cognitive and motor processes.

In an analysis of over 250 transcriptions of improvised solos recorded by the alto saxophonist Charlie Parker, Owens (1974) documented an entire system of musical patterns that Parker used in various forms throughout his solos. Finding repeated and similar patterns manually, Owens analyzed how the patterns are related to the chord progression and the key. He also outlined how some patterns are derived from others and suggested an entire system of pattern relations. Owens later extended this method of analysis to works by other jazz artists (Owens, 1996). Similarly, Finkelman (1997) showed that the guitarist Charlie Christian uses repeated melodic figures in his solos. Importantly, the goal of these investigations was to illuminate jazz improvisation using music theory. As these analyses were done by hand, statistical data about pattern use was not reported. In another antecedent to the current work, Kenny (1999) analyzed both intervallic and chordal relationships of patterns in nine transcribed solos by jazz pianist Bill Evans. He segmented the solos according to the harmonic progression so each pattern was defined by the length of the underlying chord. Repeated melodic figures spanning more than one chord were therefore not identified. In the current study, each note is considered a possible starting point for a pattern, eliminating the issue of segmentation.

Weisberg et al. (2004) used a computer algorithm to identify melodic patterns in transcriptions of 11 solos played by three jazz artists over the same chord progression. However, the authors analyzed only the melodic intervals between notes, disregarding rhythm and duration. By detecting recurrences of interval patterns, the authors were able to report pattern frequencies within and among the solos they analyzed. Weisberg and his colleagues concluded approximately 90% of the notes in Parker's solos were part of the almost 3,400 patterns identified. They also found that while shorter patterns were more common, longer patterns of up to 25 intervals did exist. It is possible that the extent of pattern use found in this research was inflated because the improvisations analyzed were all played over the same chord progression. As Johnson-Laird (2002) points out, the underlying chords serve as a constraint that may lead the improviser to use the same melodic material independent of whether this is created by inserting patterns or using procedures. It is also possible Weisberg et al. exaggerated the use of patterns as rests were ignored.

The current study extends the research by Weisberg et al. (2004) by accounting for rhythm as well as pitch parameters, and by analyzing a much larger corpus. One key concept in the current research is a strict objective definition of a musical pattern in which only interval sequences occurring twice or more is considered a pattern. For example, a melodic figure consisting of an ascending major triad followed by two descending scale steps (+4, +3, -2, -1) is only considered a pattern if that exact interval structure occurs at least twice in the corpus. A similar principle was used for rhythmic patterns where the onset times in beats were considered. Interval patterns of all lengths are reported with additional measurements reported for patterns of five notes containing four intervals (hereafter referred to as 4-interval patterns). Importantly, only interval patterns that did not include rests or notes over four beats in length are reported. All notes were considered as possible starting points for a pattern. This allowed for the reporting of the percentage of notes in the corpus that start a pattern, giving a clear indication of pattern use.

The original corpus was compared to two artificially created "control" corpora. To assure that the patterns found in the current corpus did not occur by chance, all the intervals were randomly shuffled and then analyzed. This control corpus therefore contained exactly the same interval distribution as the original corpus. In a second control analysis, the chords underlying each improvisation in the original corpus were entered into the computer program Impro-Visor (Keller, 2012) to create an alternate improvisation corpus. This program uses probabilistic grammars based on the chords and contour rules to create an improvised output (Gillick, Keller, & Tang, 2010; Keller & Morrison, 2007).

In line with the existing literature on patterns in jazz, motor sequences, and statistical pattern learning, it was predicted that examining a large corpus of improvisations from one artist would reveal extensive evidence of pattern use. If the frequency of patterns found in the original corpus was shown to significantly exceed the amount in the two control corpora, this would lend support to Pressing's (1988) theory that improvisers rely on learned patterns to generate new melodies. On the other hand, should the frequency of patterns be similar between the three corpora, then Johnson-Laird's (2002) focus on procedures would be validated by the current analysis.

Method

CORPUS

Transcriptions of 48 improvisations by Charlie Parker were included in the corpus for analysis. These solos were part of a collection transcribed by Peter Sprague (Sprague, 1988). The MIDI files were generously made available to the researcher. In addition, Parker's solo on Donna Lee, recorded on 5/8/1947, was transcribed by the researcher and analyzed. The solos represent a cross-section of Parker's improvised output from 1946-1954. Parker started recording under his own name in 1945 and made his final documented recordings in late 1954 (Jepsen, 1968; Togashi, Matsubayashi, & Hatta, 2012). A list of the solos analyzed appears in Appendix A.

In order to confirm the accuracy of the transcriptions, the researcher independently transcribed five solos without reference to the transcriptions done by Peter Sprague. To calculate an error rate, agreement versus non-agreement between the researcher's transcriptions and Peter Sprague's versions were compared. The pitch error rate was 5.7% and the rhythm error rate was 6.9%. Due to the small error rate, the corpus was considered suitable for analysis. Some discrepancies are unavoidable as transcriptions of improvised solos include some measure of interpretation. Certain notes may be ambiguous due to the quality of the recording or the phrasing of the soloists. Although the recordings have been digitally remastered, they still suffer from degraded audio quality that is a product of the recording technology during the 1940s and 1950s. Furthermore, Parker's solos include 'ghosted' notes in which he fingered the note but limited airstream. The exact pitches of those notes

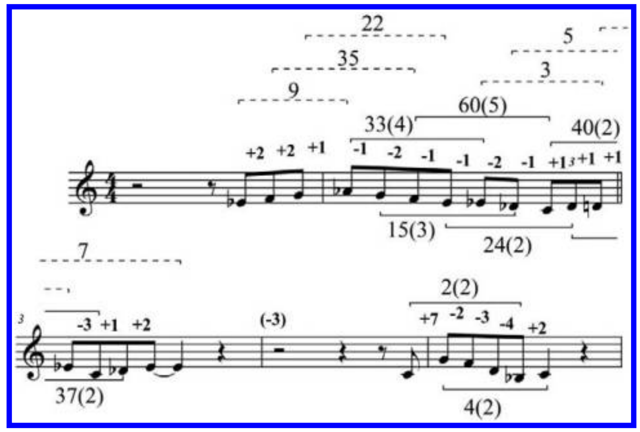

FIGURE 1. The first five measures of Parker's solo on Donna Lee. The numbers just above the notes refer to the interval. The numbers above all brackets are the number of times the pattern occurs in the corpus. The numbers in parenthesis are the number of times the pattern occurs in this solo. Perforated brackets outline interval structures that occur twice or more times in the corpus (patterns) but only once in the Donna Lee solo. Solid brackets outline patterns that are repeated both within this solo and in the corpus.

are indistinct on recordings though they have a specific rhythmic value.

The MIDI files of each solo were imported in alphabetical order into one master MIDI file in Finale, a standard music notation software package; an empty measure was inserted between each solo, and the "swing" function was added to allow for later rhythmic analysis. The master midi file was imported into the Matlab computer environment using a modified version of the MIDI Toolbox for Matlab (Smit, n.d.). This procedure converts the MIDI data into a table in which each note is listed with seven essential characteristics. These characteristics are: onset (beats), duration (beats), MIDI channel, MIDI Pitch, velocity, onset (seconds), and duration (seconds). The first 22 notes of the solo on Donna Lee are listed in Appendix B in this format.

ANALYSIS PROCEDURE

In order to analyze the corpus for the existence of relative pitch patterns a procedure was developed based on the methodology used by Weisberg et al. (2004). First the column representing absolute MIDI pitch data was converted into intervals (see Appendix B, column 4). The first interval between MIDI notes 63 (Eb4) and 65 (F4) is +2. The string of intervals in the opening notes of Parker's solo on Donna Lee therefore is +2, +2, +1, -1, -2, -1, etc. (see Figure 1). First the use of 5-note patterns was investigated by searching for 4-interval patterns with an algorithm within Matlab that did the following: Starting with the first four intervals (+2, +2, +1, -1), the program looked for additional occurrences of this interval sequence in the corpus. The result represents the number of times the interval pattern +2, +2, +1, -1 occurs in the corpus (in this case nine). The program then went on to the interval pattern starting on the following note, F4 (+2, +1, -1, -2), and looked for the number of occurrences of this pattern. Using this procedure, the number of patterns occurring on each note position can be reported. Figure 1 shows the number of times each pattern occurs in the corpus and the number of times each pattern occurs within the solo (showed in parenthesis). Patterns that only occur in the corpus are noted with perforated brackets. For example, the first note position in Figure 1, Eb, starts an interval pattern that occurs nine times in the corpus, the second note position, F, starts a pattern that occurs 35 times and so on.

Though rhythm was generally ignored in the search for pitch patterns, patterns stretching over extended rests or long notes were disregarded. In order to overcome the aforementioned limitations of Weisberg et al. (2004) regarding rests, the algorithm excluded interval patterns containing more than four beats of rest or notes longer than four beats. For example, in Figure 1, the only 4-interval patterns analyzed from the second phrase was the +7, -2, -3, -4 and the -2, -3, -4, +2 patterns. No patterns included the interval (-3) between the last note in the first phrase (Eb) and the first note of the second phrase (C). This same procedure used for 4-interval pattern analysis was repeated with pattern lengths from three intervals (four notes) to 49 intervals (50 notes), the longest identified pattern in the current corpus.

A similar procedure was used to analyze rhythm patterns. For this analysis, the first column of onset times measured in beats (see Appendix B) was used, as this measurement is independent of the tempo of each solo. Though note onsets ignore note lengths the measurement captures the most essential rhythmic feature of the music (Povel, 1984). As above, an algorithm was created that subtracted the onset beat of each note from the onset beat of the preceding note. The first couple of notes in the example listed in Figure 1 and Appendix B therefore can be listed as 0.375, 0.625, 0.375. Notice that the different lengths assigned to off-beat and on-beat eighth notes is due to the "swing value" set in Finale prior to exporting the MIDI file (see above). This allowed differentiation between identical eighth note rhythms starting on and off the beat.

Two separate tests were done to address whether the patterns found in the current corpus appeared by chance. First, the intervals in the current corpus were shuffled randomly prior to analysis using a special function in Matlab. In a second test, the chord progression underpinning the actual solos in the corpus was entered into the computer program Impro-Visor. The program was then used to generate artificial improvisations using the built-in "Parker grammar" setting. The shuffled and artificially generated corpora were analyzed for patterns using exactly the same procedures that were used for the actual corpus.

Results

The percentage of notes within the corpus that start interval and rhythm patterns of various lengths were investigated using a specially designed algorithm within Matlab. In total, 82.6% of the eligible notes in the corpus began a 4-interval pattern occurring at least twice in the corpus. Eligible notes are notes that start a pattern that is not interrupted by a four beat rest or long note (see methods). The percentage of notes beginning 4-interval patterns that occur at least three times was 73.9%. Half the notes in the corpus started a 4-interval pattern that occurs nine times or more. Ten percent of the notes started a 4-interval pattern that occurs 70 times or more. Figure 2 lists the percentage of notes that start a 4-interval pattern as a function of the number of times the pattern occurs in the corpus. The mean number of times the 4-interval pattern on each note position is repeated in the solos analyzed was 26.3. In total 99.3% of all the notes in the corpus were part of patterns consisting of three or more intervals that occurred two or more times.

When limiting pattern identifications to sequences containing the same four intervals and rhythms, 57.6% of notes start patterns that occur at least twice in the corpus. It is 48.2% with three or more occurrences. Figure 3 lists the percentage of notes that start a 4-interval and rhythm pattern as a function of the number of times the pattern occurs in the corpus.

Investigating various lengths of interval patterns, the mean number of occurrences of a 3-interval pattern is 83.6 for each eligible note in the corpus. This number decreases quickly as the length of the pattern is increased as seen in Figure 4. However, the high number of three-interval patterns is probably due to the fewer number of possible patterns with only four notes. Figure 5 lists the number of unique pattern combinations found in the corpus by length for up to 19-interval patterns. There were more unique combinations of 5-interval patterns than any other pattern length.

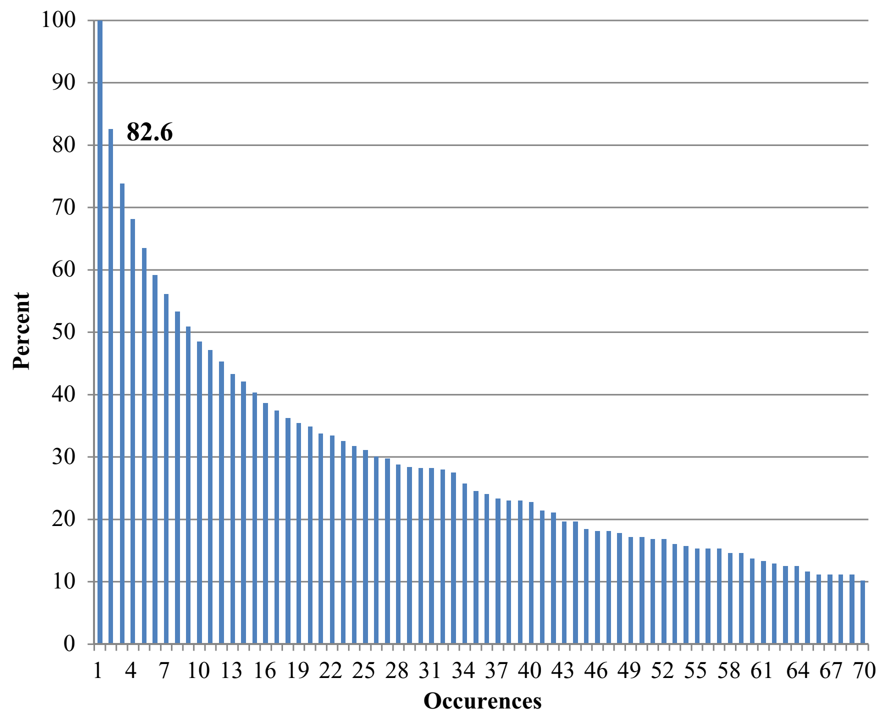

FIGURE 2. The percentage of notes that start a 4-interval pattern as a function of the number of times the pattern occurs in the corpus.

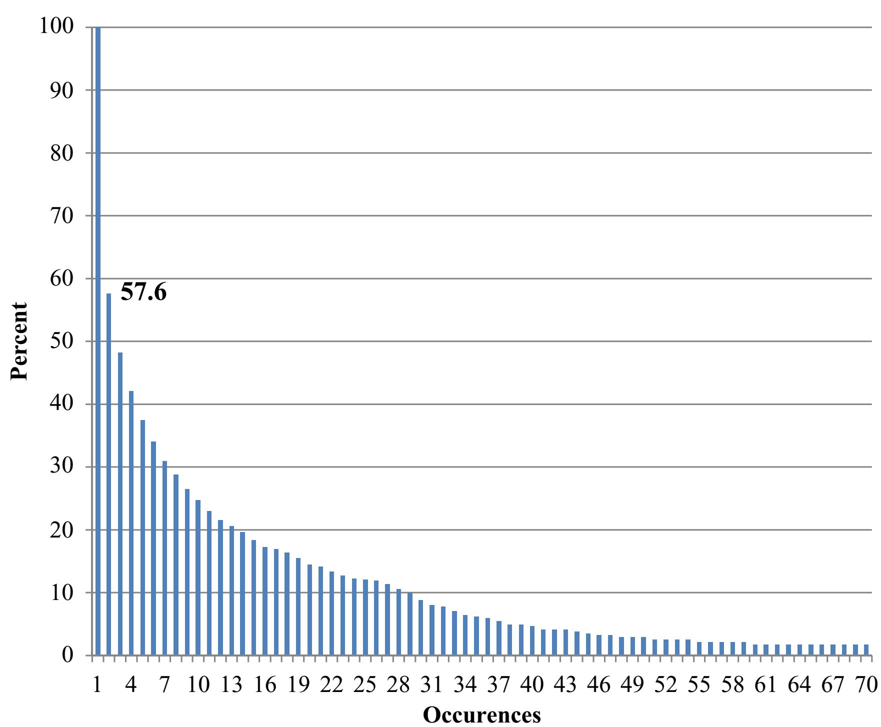

FIGURE 3. The percentage of notes that start a 4-interval and rhythm pattern as a function of the number of times the pattern occurs in the corpus.

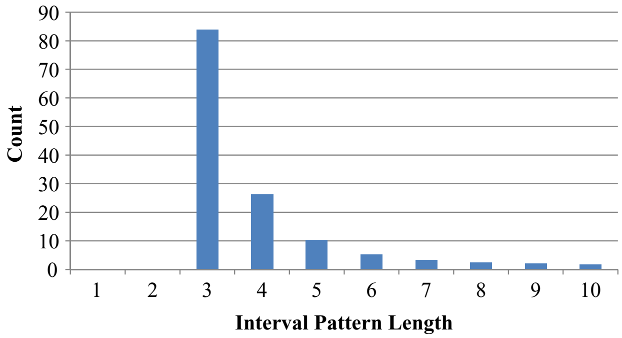

FIGURE 4. The mean number of occurrences on each note position as a function of pattern length. Data shown for patterns up to 10 intervals long.

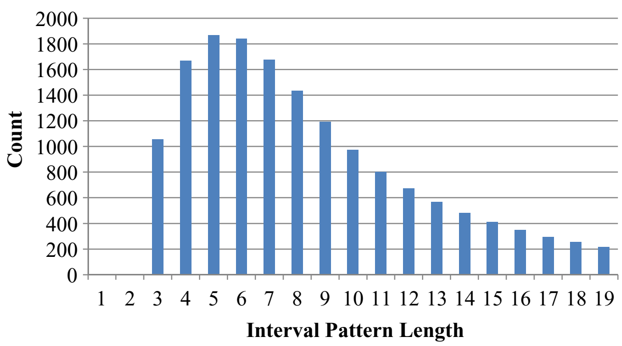

FIGURE 5. The number of unique patterns in the corpus as a function of interval pattern length. Data shown for patterns up to 19 intervals long.

The longest interval pattern identified was two occurrences of a 49-interval (50-note) pattern. It follows from the definition of patterns as outlined above, that any 5-interval pattern will contain two overlapping 4-interval patterns. For example, the pattern in the second phrase in Figure 1 (+7, -2, -3, -4, +2) contains both the +7, -2, -3, -4 and the -2, -3, -4, +2 pattern. Therefore, any 49-interval pattern will contain two 48-interval patterns, three 47-interval patterns, etc.

To investigate the most common length of patterns, an algorithm was designed that disregarded shorter patterns contained within longer patterns. The algorithm segmented the entire corpus into the longest possible pattern as follows. Initially, the algorithm looked for the longest patterns identified in the earlier analysis, the 49-interval pattern. After identifying that pattern, the algorithm "erased" the notes within that pattern and ran the analysis looking for 48-interval patterns. The corpus contains no 48-interval patterns that are not contained within a longer pattern. The algorithm then continued to identify 47-interval patterns all the way down to three-interval patterns. In this analysis, these three-interval patterns were not part of any longer pattern. Using this algorithm, the mean length of interval patterns was calculated to be 7.3 intervals.

Anecdotal evidence suggests that the patterns brought to light by the segmentation algorithm often represent patterns that occur within the same solo. In Figure 6, the first 22 measures of Parker's solo on Donna Lee are segmented two ways. The perforated brackets outline the longest pattern as compared to the solo only, the solid brackets as compared to the corpus. Notice that in several instances the two types of brackets align.

To investigate whether longer patterns mostly occur within the same solo or in many solos throughout the corpus, the non-overlapping patterns identified with the algorithm mentioned above were further analyzed for their placement. Patterns longer than 15-intervals that are not contained within other patterns are listed in Figure 7. Note that the four longest patterns in the corpus (30-, 32-, 34-, and 49-intervals long) were played in the same solo. However, the majority of patterns between 15 and 29 intervals long occurred in multiple solos. Of the 98 unique non-overlapping patterns over 15-intervals long identified in the corpus, 38 occurred in only one solo while 60 were present in multiple solos.

An analysis of the most frequently occurring 4-interval patterns showed the most common pattern was the interval structure representing the five notes of a descending major scale starting on the 5^{th} scale degree and descending to the root. However, as the tonal context was not analyzed in this study, the underlying chord cannot be inferred. Appendix C shows the 30 most frequently occurring 4-interval patterns. Notice that the most common patterns are descending combinations of half and whole steps. Common ascending patterns often include triad arpeggios. In reality, many of these 4-interval patterns were part of the same longer pattern as this list includes overlapping patterns. For example, -2, -1, -2, -2 (184 occurrences) and -1, -2, -2, -1 (139 occurrences) is often combined in a five-interval pattern, -2, -1, -2, -2, -1 (60 occurrences); the interval structure of a descending major scale from the 5^{th} to the seventh scale degree. A full melodic and harmonic analysis of Parker's vocabulary has been covered elsewhere (Owens, 1974) and is beyond the scope of this paper.

The context in which patterns appear is of interest as it relates to the discussion of how motor patterns are linked together. Figure 8 shows four instances of the appearance of a 10-interval pattern that occurs 22 times in the corpus. The different examples of the pattern illustrate how the same pattern can begin a phrase or be linked to additional notes on either side. It also shows

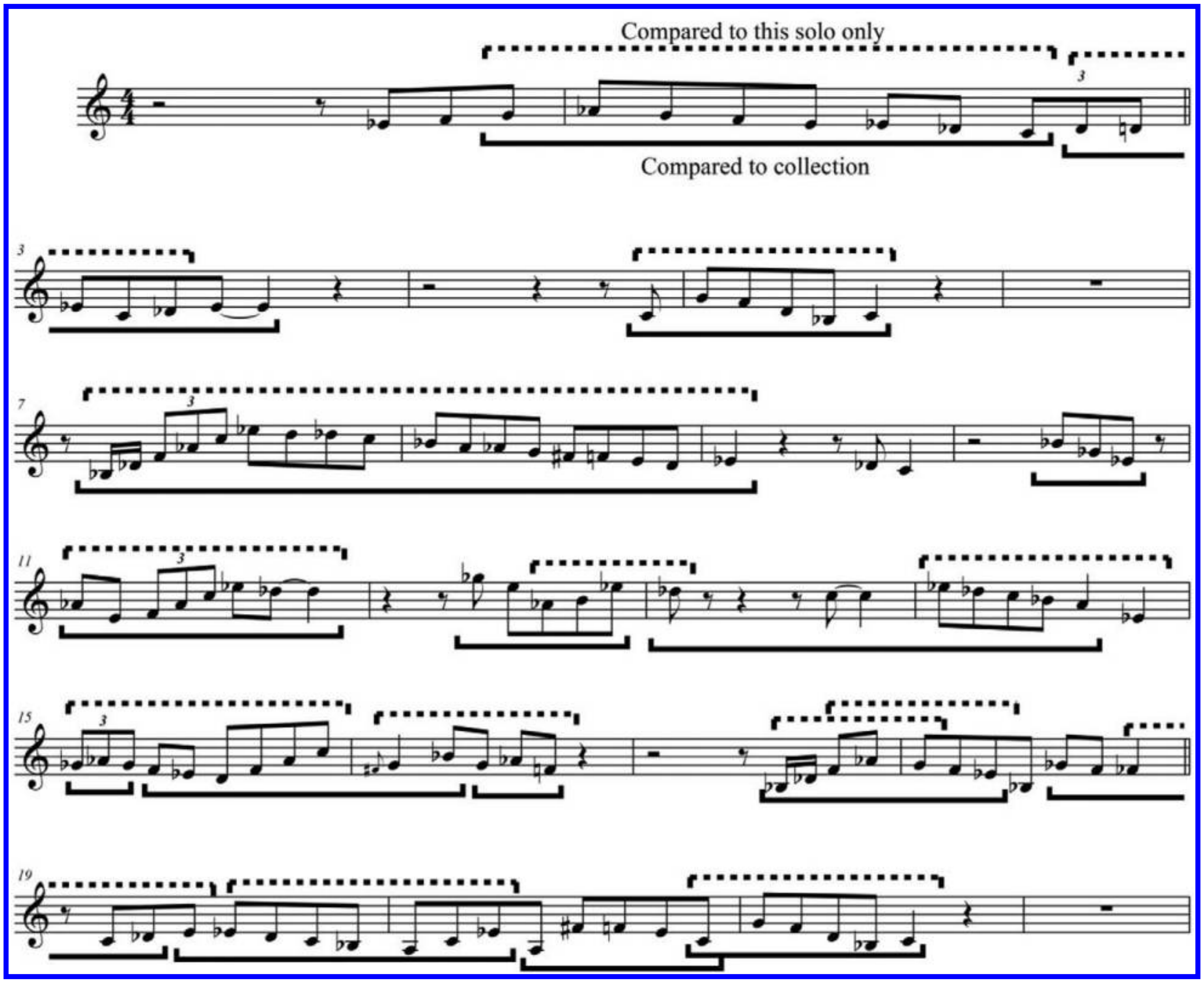

FIGURE 6. The first 22 measures of Parker's solo on Donna Lee segmented into the longest possible pattern as compared to the solo (perforated bracket) and the corpus (solid bracket).

how the pattern may be performed with various rhythms and in two different keys.

As mentioned earlier, it has been suggested that the presence of patterns in improvised solos can be attributed to improvisers following tonal rules dictated by the chords. To investigate this issue, two different analyses were undertaken. First, to examine if the patterns in the current corpus simply appeared by chance due to the large number of intervals analyzed, 4-interval patterns were extracted after the intervals were shuffled randomly. The interval distribution was therefore identical in the original and randomized versions. Indeed, 60.1% of the notes started a 4-interval pattern in the randomized version if a pattern were defined as occurring two or more times. However, for higher number of occurrences, the numbers differed substantially from the actual patterns found in the corpus (compare Figure 9 to Figure 2). For example, if a pattern is defined as a 4-interval structure occurring 10 times or more, only 6.1% of the notes in the randomized version started a pattern as compared to 48.5% in the actual corpus.

A second analysis was conducted on a computer-generated corpus to investigate whether the patterns simply appeared because the improvisations followed tonal rules and jazz convention. According to the programmers, the computer program Impro-Visor uses algorithms to generate melodic solos based on a given chord progression (Gillick et al., 2010; Keller & Morrison, 2007). Importantly, the program extracts algorithms

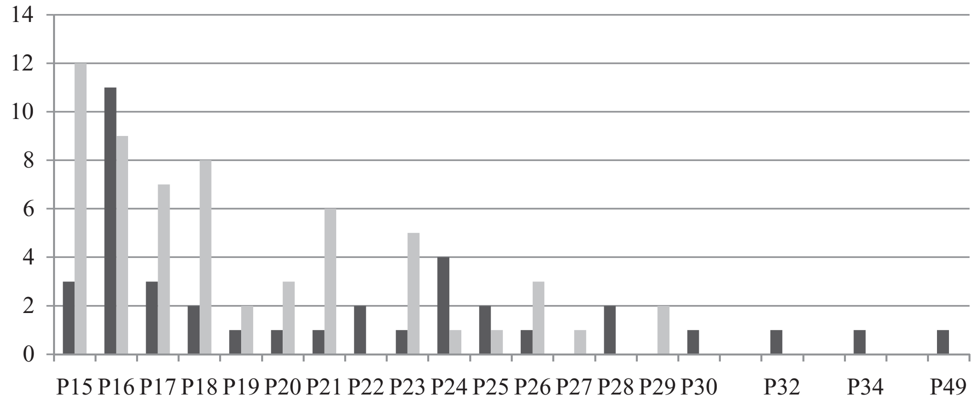

FIGURE 7. All non-overlapping patterns of 15-interval lengths or more are shown. The dark columns represent the number of different patterns that only occur within one solo. The lighter columns represent the number of patterns that occur in more than one solo.

FIGURE 8. Examples of a 10-interval (11-note) pattern denoted with brackets shown in context. Excerpts from Parker's solo on *Bird of Paradise, I Get a Kick Out of You*, and *She Rote*.

based on actual solos but does not store and reuse specific patterns. Only 52.2% of the notes in the computer-generated corpus started a 4-interval pattern that occurred twice or more (compared to 60.1% for the shuffled corpus and 82.6% for the actual corpus). For patterns occurring 10 times or more, only 7.2% of the notes in the artificial corpus began a pattern compared to 48.5% in the Parker corpus. Figure 10 lists the percentage of notes that start a 4-interval pattern as a function of the number of times the pattern occurs in the artificial corpus. In the artificial corpus, the longest interval pattern identified was only 10 intervals long as

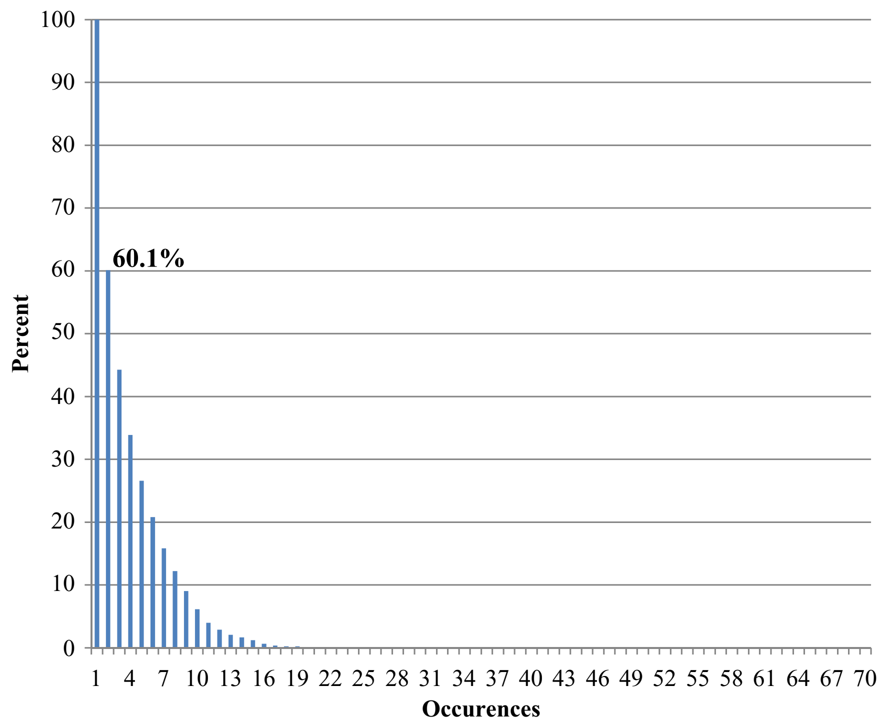

FIGURE 9. The percentage of notes that start a 4-interval pattern as a function of the number of times the pattern occurs in the corpus with intervals shuffled randomly.

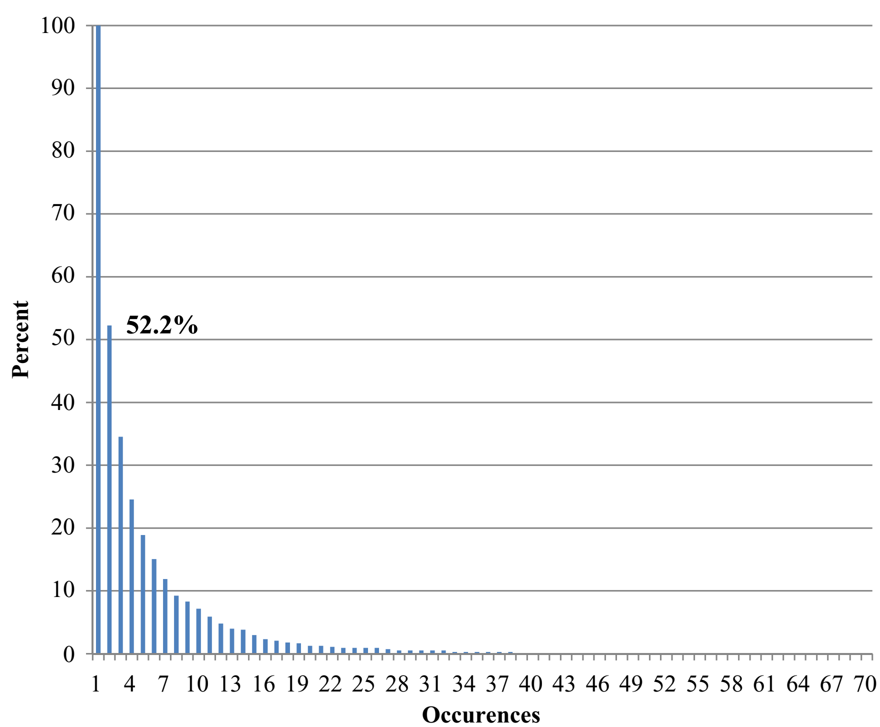

FIGURE 10. The percentage of notes that start a 4-interval pattern as a function of the number of times the pattern occurs in the computer-generated corpus. The corpus was generated using the chord progressions from the Parker corpus.

compared with 49 intervals in the actual corpus. Only one unique version of this pattern occurred in the artificial corpus compared to 971 unique 10-interval patterns in the actual corpus.

Discussion

The current study investigated the use of patterns in a large corpus of Charlie Parker solos. Results showed 82.6% of the notes in the current corpus started a 4-interval (5-note) pattern and that the mean number of times these patterns were repeated in the corpus was 26.3. In addition, 57.6% of the notes in the corpus began patterns in which both the interval structure and the rhythm was identical. This is the first time the number of note positions that start a pattern has been reported in a computer guided analysis of a large corpus of improvised solos. This measurement accurately reflects the use of patterns in improvisations and could be used to compare artists, styles, and periods.

It should be noted that the pattern-detection algorithm used here was not designed to detect patterns repeated with slight variations. This type of analysis requires that the identified patterns are compared and categorized according to predetermined rules. This necessarily involves a subjective assessment of the degree of similarity that is intentionally avoided in the current work. Previous research used music theoretical concepts including the underlying chord structure to segment solos into patterns and divide related patterns into categories (Finkelman, 1997; Kenny, 1999, Owens, 1974). Future research could augment the current computer aided paradigm by coding each note's relationship to the underlying chord and thereby categorize similar patterns.

Parker's ubiquitous use of patterns appears to lend support to Pressing's theoretical model for improvisation in which learned auditory and motor patterns play a central role in the process of generating music (Pressing, 1988). In contrast, Johnson-Laird's (2002) notion that motor and auditory patterns are not stored and reused during improvisation appears in conflict with the current findings. Specifically, an analysis of placements of longer (over 15 intervals) non-overlapping patterns showed that 61% of those patterns occurred in more than one improvisation. In total, nearly every single note (99.3%) in the current corpus was part of a pattern. Notes not included in patterns usually occurred right before or immediately following long gaps. Recall that the detection algorithm used here did not identify patterns across rest or note values over four beats in length. Patterns are not simply a feature of the current corpus, they virtually are the corpus. This is not to say that procedures for creating improvised lines are not active in addition to the insertion of learned patterns. However, the current results suggest that learned patterns are essential to the improvisational process.

One can argue that the analysis of an existing corpus may not lead to an accurate interpretation of how the notes are created. However, new experimental research mentioned previously supports the interpretation above (Goldman, 2012). According to Goldman, improvisers "can execute motor patterns according to what they 'hear'" (p. 367) when they play in familiar keys. However, forced to play in an unfamiliar key they employ knowledge "which can be used to generate new, unrehearsed motor patterns" (p. 367). As a highly accomplished performer and recording artist in the jazz idiom, Charlie Parker would have had ample opportunity to acquire a repository of patterns appropriate to the keys and chord progressions of the tunes represented in the corpus. Lacking such a resource, a novice improviser would be more likely to rely on learned procedures when soloing. Similarly, experienced performers soloing over an unfamiliar chord progression may depend upon learned procedures to a greater extent than they would in more familiar tunes.

To address concerns that the patterns found in the current corpus could have appeared by chance, two other analyses were undertaken. First, the intervals in the original corpus were shuffled randomly and analyzed for the existence of patterns. This analysis showed far fewer patterns than the original corpus (compare Figure 9 to Figure 2). Secondly, a corpus of computer-generated improvisations was created using a rule-based algorithm on exactly the same underlying chord progressions as the original corpus. This alternate corpus did not show a similar structure of 4-interval patterns as compared to the Parker corpus (compare Figure 10 to Figure 2). The patterns in the alternate corpus were also much shorter and fewer in number than the patterns in the Parker corpus. It is, however, possible that this was due to the algorithm used by the computer program not accurately reflecting the style of Charlie Parker. Future improvements to this or similar computer programs may result in artificially created tonal improvisations with realistic pattern structures. Specifically, such an algorithm should produce improvisations that include a large number of scale fragments and arpeggios similar to the most common patterns in the current corpus (see Appendix C). This would strengthen the case that patterns may appear even if the method of creation is based on procedures. Alternatively, it may be possible to create a computer program that models the creative process of jazz improvisers by using stored patterns. The current author is currently collaborating with researchers in mathematics and computer science to create such a program (Norgaard, Spencer, & Montiel, 2013). Future research in computer modeling could investigate whether pattern use or procedures best model the thinking of artist-level jazz improvisers.

The use of patterns by artist-level jazz improvisers playing saxophone (Owens, 1974), guitar (Finkelman, 1997), and piano (Goldman, 2012; Kenny, 1999) aligns with motor research showing that longer action sequences may be a concatenation of smaller submovements (Park & Shea, 2005). The submovements are stored general motor programs (GMPs) that are effector independent and contain relative timing, sequence, and force information (Shea & Wulf, 2005). Through a recall schema this abstract stored information is converted to the actual situation and linked with other GMPs. Similarly, musical improvisation may involve a process in which the improviser uses stored GMPs that are linked together (Pressing, 1988). These stored GMPs may contain both motor and auditory information that has been linked through practice and possibly stored in the same brain region (Baumann et al., 2007; Lahav et al., 2007). Though the data from the current study aligns with the general concept of motor research, it is gathered from a particular type of melodic improvisation within the jazz tradition. Inferences to other types of improvisation should be made with caution.

It is possible that advanced improvisers use learned procedures for concatenating patterns during improvisation. This view is consistent with the data that shows that some shorter patterns are used very often (about 10% of the 4-interval patterns occur 70 times or more) while others occur rarely (see Figure 2). The patterns that occur rarely may simply appear by chance as musicians use learned procedures to link stored patterns. These procedures would be based on the underlying harmony and meter, and based on tonal rules (Berliner, 1994; Jackendoff & Lerdahl, 2006; Johnson-Laird, 2002). It may be that Johnson-Laird's focus on learned procedures described earlier is adaptable to the notion of stored patterns if those procedures describe how the patterns are linked. It is also possible that the concatenation of patterns is partly based on the movements themselves (Novembre & Keller, 2011). Indeed, Novembre and Keller suggested that pianists may develop a movement grammar through extensive practice that is linked to syntactical musical rules. The current research showed the existence of a large number of patterns in the improvisations of Charlie Parker but did not describe how those patterns are linked. Future research could investigate this issue by calculating the probability that one pattern is followed by another and comparing this information to related tonal rules.

The ubiquity of patterns in the current research suggests an underlying learning mechanism that favors pattern learning (Gobet et al., 2001). Johnson-Laird (2002) suggests that from a computational standpoint, patterns would have to first be created in order to appear in an improvisation. Obviously, advanced improvisers can create material simply based on tonal rules. However, interviews with jazz musicians have documented the use of patterns as a deliberate learning tool (Berliner, 1994) and many traditional jazz pedagogical materials include patterns (Baker, 1988; Coker, 1980). A pattern is initially learned as both an auditory image and corresponding motor realization (Drost et al., 2005). The latter will be linked to the key in which the pattern is initially learned. Through practice and experience the same auditory image may be linked to many motor representations necessary to execute the pattern in various keys. It is possible that patterns are also learned serendipitously while listening and practicing. Indeed research in language learning suggests that infants use statistical learning processes to identify word boundaries and that this process is domain independent (Saffran et al., 1996, 1999). Similarly, developing jazz musicians may use statistical processes during listening to identify common musical patterns. Furthermore, it appears this learning process happens incidentally during language acquisition (Saffran et al., 1997) suggesting that musicians may store patterns subconsciously without attending to the musical source. The improviser may be able to automatically link these auditory patterns to their related motor representation through a strong learned auditory-motor association (Keller & Koch, 2008). Novice improvisers may also store linked auditory and motor patterns by simply monitoring their own playing and identifying salient patterns consciously and subconsciously.

In summary, the current research suggests that the use of patterns is central to musical improvisation in agreement with the theoretical framework suggested by Pressing (1988). This is based both on a statistical analysis of a large corpus of improvised solos by Charlie Parker and on a literature review of related motor and language research. Though procedures based on tonal rules may guide improvisations on novel chord sequences, this process is likely superseded by the insertion of patterns as the context becomes familiar. These patterns may be acquired through deliberate practice or incidentally through statistical learning processes and concatenated during performance to produce the final improvised output. These results may have implications for other domains in which human action sequences are created in real time.

Author note

I would like to thank Philip Johnson-Laird and James Fidlon for helpful comments and suggestions on an earlier version of this manuscript. I am also grateful to two anonymous reviewers for their constructive criticism.

Correspondence concerning this article should be addressed to Martin Norgaard, Georgia State University, School of Music, PO Box 4097, Atlanta, GA 30302-4097. E-mail: mnorgaard@gsu.edu

References

Baker, D. N. (1988). *Jazz improvisation* (2nd ed.). Van Nuys, CA: Alfred.

Baumann, S., Koeneke, S., Schmidt, C. F., Meyer, M., Lutz, K., & Jancke, L. (2007). A network for audio-motor coordination in skilled pianists and non-musicians. *Brain Research*, *1161*, 65-78.

Berliner, P. F. (1994). *Thinking in jazz*. Chicago, IL: University Of Chicago Press.

Chen, J. L., Rae, C., & Watkins, K. E. (2012). Learning to play a melody: An fMRI study examining the formation of auditory-motor associations. *NeuroImage, 59,* 1200-1208.

Coker, J. (1980). *Jerry Coker's complete method for improvisation: For all instruments*. Van Nuys, CA: Alfred.

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

Dietrich, A., & Kanso, R. (2010). A review of EEG, ERP, and neuroimaging studies of creativity and insight. *Psychological Bulletin*, *136*, 822-848.

Drost, U. C., Rieger, M., Gunter, T. C., & Prinz, W. (2005). When hearing turns into playing: Movement induction by auditory stimuli in pianists. *Quarterly Journal of Experimental Psychology*, *58A*, 1376-1389.

Finkelman, J. (1997). Charlie Christian and the role of formulas in jazz improvisation. *Jazzforschung/Jazz Research*, *29*, 159-188.

Gillick, J., Tang, K., & Keller, R. M. (2010). Machine learning of jazz grammars. *Computer Music Journal*, *34(3)*, 56-66.

Gobet, F., Lane, P. C. R., Croker, S., Cheng, P. C., Jones, G., Oliver, I., & Pine, J. M. (2001). Chunking mechanisms in human learning. *Trends in Cognitive Sciences*, *5*, 236-243.

Goldman, A. (2012). What does one know when one knows how to improvise? In E. Cambouropoulos, C. Tsougras, P. Mavromatis, & K. Pastiadis (Eds.), *Proceedings of the 12th International Conference on Music Perception and Cognition and the 8th Triennial Conference of the European Society for the Cognitive Sciences of Music* (pp. 360-368). Thessaloniki, Greece: ICMPC. Retrieved from http://icmpc-escom2012.web.auth.gr/?q=node/67

Goldman, A. (2013). Towards a cognitive-scientific research program for improvisation: Theory and an experiment. *Psychomusicology: Music, Mind, and Brain, 23*, 210-221.

Hay, J. F., Pelucchi, B., Estes, K. G., & Saffran, J. R. (2011). Linking sounds to meanings: Infant statistical learning in a natural language. *Cognitive Psychology*, *63*, 93-106.

Jackendoff, R., & Lerdahl, F. (2006). The capacity for music: What is it, and what's special about it? *Cognition*, *100*, 33-72.

Jepsen, J. G. (1968). *A discography of Charlie Parker.* Copenhagen: Karl Emil Knudsen.

Johnson-Laird, P. N. (2002). How jazz musicians improvise. *Music Perception*, *19*, 415-442.

Keller, P. E., & Koch, I. (2008). Action planning in sequential skills: Relations to music performance. *Quarterly Journal of Experimental Psychology*, *61*, 275-291.

Keller, R. (2012). Impro-Visor [Computer software]. Retrieved from http://sourceforge.net/projects/impro-visor/

Keller, R., & Morrison, D. (2007). A grammatical approach to automatic improvisation. In C. Spyridis, A. Georgaki, G. Kouroupetroglou, & C. Anagnostopoulou (Eds.), *Proceedings of the 4th Sound and Music Computing Conference* (pp. 330-337). Lefkada, Greece: National and Kapodistrian University of Athens. Retrieved from http://www.smc-conference.net/smc07/SMC07 Proceedings/SMC07 Paper 55.pdf

Kenny, B. (1999). Structure in jazz improvisation: A formulaic analysis of the improvisations of Bill Evans. *Annual Review of Jazz Studies 10,* 163-194.

Lahav, A., Saltzman, E., & Schlaug, G. (2007). Action representation of sound: Audiomotor recognition network while listening to newly acquired actions. *Learning*, *27*, 308-314.

Lew-Williams, C., & Saffran, J. R. (2012). All words are not created equal: Expectations about word length guide infant statistical learning. *Cognition*, *122*, 241-246.

Novembre, G., & Keller, P. E. (2011). A grammar of action generates predictions in skilled musicians. *Consciousness and Cognition*, *20*, 1232-1243.

Norgaard, M. (2011). Descriptions of improvisational thinking by artist-level jazz musicians. *Journal of Research in Music Education*, *59*, 109-127.

Norgaard, M., Spencer, J., & Montiel, M. (2013). Testing cognitive theories by creating a pattern-based probabilistic algorithm for melody and rhythm in jazz improvisation. *Psychomusicology: Music, Mind, and Brain*, *23*, 243-254.

Owens, T. (1974). *Charlie Parker: Techniques of improvisation* (Unpublished doctoral dissertation). University of California, Los Angeles.

Owens, T. (1996). *Bebop: The music and its players*. New York: Oxford University Press.

Park, J.-H., & Shea, C. H. (2005). Sequence learning: Response structure and effector transfer. *The Quarterly Journal of Experimental Psychology, 58*, 387-419.

Povel, D. (1984). A theoretical framework for rhythm perception. *Psychological Research*, *45*, 315-337.

Pressing, J. (1988). Improvisation: Methods and model. In J. A. Sloboda (Ed.), *Generative processes in music* (pp. 129-178). Oxford, UK: Oxford University Press.

Saffran, J. R. (2003). Statistical language learning: Mechanisms and constraints. *Current Directions in Psychological Science*, *12*, 110-114.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*, 1926-1928.

Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, *70*, 27-52.

Saffran, J. R., Newport, E. L., Aslin, R. N., Tunick, R. A., & Barrueco, S. (1997). Incidental language learning: Listening (and learning) out of the corner of your ear. *Psychological Science*, *8*, 101-105.

Saffran, J. R., Pollak, S. D., Seibel, R. L., & Shkolnik, A. (2007). Dog is a dog is a dog: Infant rule learning is not specific to language. *Cognition*, *105*, 669-680.

Schmidt, R. A. (1975). A schema theory of discrete motor skill learning. *Psychological Review*, *82*, 225-260.

Schmidt, R. A., & Lee, T. D. (2011). *Motor control and learning: A behavioral emphasis. Champaign, IL: Human Kinetics* (5th ed.). Champaign, IL: Human Kinetics.

Shea, C. H., & Wulf, G. (2005). Schema theory: A critical appraisal and reevaluation. *Journal of Motor Behavior*, *37*, 85-101.

Smit, C. (n.d.). KaraokeMidiJava [Computer software]. Retrieved from http://www.ee.columbia.edu/~csmit/matlab_midi.html

Sprague, P. (1988). *The jazz solos of Charlie Parker*. Del Mar, CA: Spragueman.

Temperley, D. (2007). *Music and probability*. Cambridge, MA: MIT Press.

Togashi, N., Matsubayashi, K., & Hatta, M. (2012). Jazz discography project. Retrieved from http://www.jazzdisco.org/charlie-parker/discography/

Weisberg, R. W., Brinkman, A. R., Folio, C. J., Dick, A. S., Fleck, J. I., Niederberger, B., & Barrett, F. (2004). Toward a cognitive analysis of creativity: Improvisation in jazz. In R. Parncutt, A. Kessler, & F. Zimmer (Eds.), *Proceedings of the Conference on Interdisciplinary Musicology* (pp. 1-9). Graz, Austria: Conference on Interdisciplinary Musicology. Retrieved from http://www.uni-graz.at/richard.parncutt/cim04/CIM04_paper_pdf/Weisberg_Brinkman_Folio_Dick_Fleck_Niederberg_Barrett_CIM04_proceedings.pdf

Appendix A

LIST OF RECORDED IMPROVISATIONS BY CHARLIE PARKER INCLUDED IN THE CORPUS

Song Title	Album or CD Title	Catalogue Number	Date
Moose the Mooche	Charlie Parker on Dial, Vol. 1	D1010-2	3/28/1946
Yardbird Suite take 2	Charlie Parker on Dial, Vol. 1	D1011-1	3/28/1946
Yardbird Suite take 1	Charlie Parker on Dial, Vol. 1	D1011-4	3/28/1946
Ornithology	Charlie Parker on Dial, Vol. 1	D1012-4	3/28/1946
Bird's Nest	Charlie Parker on Dial, Vol. 2	D1053-C	2/19/1947
Stupendous take 1	The Very Best of Bird	D1074-A?	2/26/1947
Stupendous take 2	Charlie Parker Immortal Sessions, Vol. 8	D1074-B	2/26/1947
Donna Lee	Yardbird Suite disc one	S3420-3	5/8/1947
Dexterity	Charlie Parker on Dial, Vol. 4	D1101-B	10/28/1947
Bongo Bop take 2	Charlie Parker on Dial, Vol. 4	D1102-A	10/28/1947
Bongo Bop take 1	Charlie Parker on Dial, Vol. 4	D1102-B	10/28/1947
Bird of Paradise	Charlie Parker on Dial, Vol. 4	D1105-C	10/28/1947
Scrapple from the Apple	Charlie Parker on Dial, Vol. 5	D1113-C	11/4/1947
Driftin' on a Reed	Charlie Parker on Dial, Vol. 6	D1151-E	12/17/1947
Bongo Beep	Charlie Parker on Dial, Vol. 6	D1154-C	12/17/1947
Crazeology	Charlie Parker on Dial, Vol. 6	D1155-D	12/17/1947
Merry-Go-Round	The Complete Charlie Parker, Vol. 5	B911-?	Aug/Sep 1948
Perhaps take 1 (original take 6)	The Complete Charlie Parker, Vol. 5	B908-6?	Aug/Sep 1948
Perhaps take 2	The Complete Charlie Parker, Vol. 5	B908-?	Aug/Sep 1948
Just Friends	Charlie Parker, The Verve Years (1948-1950)	319-5	11/30/1949
Star Eyes	The Definitive Charlie Parker, Vol. 2	371-4	March/April 1950
Bloomdido	The Definitive Charlie Parker, Vol. 2	410-4	6/6/1950

(continued)

Appendix A (continued)

Song Title	Album or CD Title	Catalogue Number	Date
An Oscar for Treadwell	The Definitive Charlie Parker, Vol. 2	411-3	6/6/1950
My Melancholy Baby	The Definitive Charlie Parker, Vol. 2	413-2	6/6/1950
Leap Frog take 1	The Definitive Charlie Parker, Vol. 2	414-4	6/6/1950
Leap Frog take 2	The Definitive Charlie Parker, Vol. 2	414-6	6/6/1950
She Rote take 1	The Definitive Charlie Parker, Vol. 4	490-3	1/17/1951
She Rote take 2	The Definitive Charlie Parker, Vol. 4	490-5	1/17/1951
K. C. Blues	The Definitive Charlie Parker, Vol. 4	491-1	1/17/1951
Un Poquito De Tu Amour	The Definitive Charlie Parker, Vol. 4	541-2	3/12/1951
Why Do I Love You? take 1	The Definitive Charlie Parker, Vol. 4	544-2	3/12/1951
Why Do I Love You? take 2	The Definitive Charlie Parker, Vol. 4	544-6	3/12/1951
Why Do I Love You? take 3	The Definitive Charlie Parker, Vol. 4	544-7	3/12/1951
Blues For Alice	The Definitive Charlie Parker, Vol. 4	609-4	8/8/1951
Si Si	The Definitive Charlie Parker, Vol. 5	610-4	8/8/1951
Swedish Schnapps	The Definitive Charlie Parker, Vol. 5	611-3	8/8/1951
Back Home Blues	The Definitive Charlie Parker, Vol. 5	612-1	8/8/1951
Kim	The Definitive Charlie Parker, Vol. 6	1120-2	12/30/1952
In the Still of the Night	The Definitive Charlie Parker, Vol. 7	1238-7	5/22/1953
Old Folks	The Definitive Charlie Parker, Vol. 7	1239-9	5/22/1953
If I Love Again	The Definitive Charlie Parker, Vol. 7	1240-9	5/22/1953
Confirmation	The Definitive Charlie Parker, Vol. 7	1249-3	8/4/1953
I Get a Kick out of You take 1	The Definitive Charlie Parker, Vol. 7	1531-2	3/11/1954
I Get a Kick out of You take 2	The Definitive Charlie Parker, Vol. 7	1531-7	3/11/1954
Just One of those Things	The Definitive Charlie Parker, Vol. 7	1532-1	3/11/1954
I've Got You Under My Skin	The Definitive Charlie Parker, Vol. 7	1534-1	3/11/1954
Love for Sale take 1	The Definitive Charlie Parker, Vol. 8	2115-4	12/10/1954
Love for Sale take 2	The Definitive Charlie Parker, Vol. 8	2115-5	12/10/1954

Appendix B

THE FIRST 22 NOTES OF PARKER'S SOLO ON DONNA LEE SHOWN AFTER CONVERSION TO THE MATLAB ENVIRONMENT

Onset (beats)	Duration (beats)	Midi channel	Midi Pitch	Velocity	Onset (seconds)	Duration (seconds)
2.625	0.375	1	63	64	0.875	0.125
3.000	0.625	1	65	64	1.000	0.208
3.625	0.375	1	67	64	1.208	0.125
4.000	0.625	1	68	64	1.333	0.208
4.625	0.375	1	67	64	1.542	0.125
5.000	0.625	1	65	64	1.667	0.208
5.625	0.375	1	64	64	1.875	0.125
6.000	0.625	1	63	64	2.000	0.208
6.625	0.375	1	61	64	2.208	0.125
7.000	0.333	1	60	64	2.333	0.111
7.333	0.334	1	61	64	2.444	0.111
7.667	0.333	1	62	64	2.556	0.111
8.000	0.625	1	63	64	2.667	0.208
8.625	0.375	1	60	64	2.875	0.125
9.000	0.625	1	61	64	3.000	0.208
9.625	1.375	1	63	64	3.208	0.458
15.625	0.375	1	60	64	5.208	0.125
16.000	0.625	1	67	64	5.333	0.208
16.625	0.375	1	65	64	5.542	0.125
17.000	0.625	1	62	64	5.667	0.208
17.625	0.375	1	58	64	5.875	0.125
18.000	1.000	1	60	64	6.000	0.333

Appendix C

THE 30 MOST COMMONLY USED 4-INTERVAL (5-NOTE) PATTERNS IN THE CORPUS

number of occurrences	Pattern			
184	-2	-1	-2	-2
171	-1	-1	-1	-1
147	-2	-1	-2	-1
139	-1	-2	-2	-1
121	-2	-2	-1	-2
117	-2	-1	3	3
95	2	-2	-1	-2
92	-1	-2	-1	3
87	3	-2	-1	-2
87	1	3	4	3
86	1	2	-2	-1
80	3	4	3	-3
70	-1	-1	-1	-2
69	3	4	3	-2
69	-1	3	3	3
65	-3	1	1	-4
64	4	3	-3	1
64	-1	-2	-2	-2
62	3	3	-1	-2
61	1	-1	-2	-2
60	-1	-1	-2	-1
59	3	-3	1	1
59	-2	-1	-9	3
57	-2	-2	-2	-1
57	-2	-2	-1	3
54	-1	-9	3	3
53	2	2	1	2
52	-1	-2	-1	-9
52	-1	-2	-1	-2
50	9	-3	1	1