SPECTROGRAPHIC AND CALLIGRAPHIC CUES IN THE IDENTIFICATION OF JAZZ SAXOPHONISTS

Fernando Benadon
University of California, Berkeley, USA

ABSTRACT

Background. Conceptually, it has been common to distinguish two aspects of music performance, what might be called *calligraphic* (i.e., pitch, rhythm, and contour) and *spectrographic* (i.e., timbre, non-quantized rhythm, and expressive gestures such as pitch bend and vibrato). Where scholars have investigated either calligraphic or spectrographic features of jazz, they have failed to examine rigorously which of the features make the performer recognizable by listeners and therefore perceptually unique.

Aims. An experiment is presented to support the hypothesis that well known jazz saxophonists can be recognized through predominantly spectrographic information.

Method. Seven volunteers participated in the study. The stimuli consisted of sixteen short audio samples drawn from post-bebop recordings. Each excerpt featured a tenor saxophonist playing between 2 and 5 notes while accompanied by a rhythm section. Subjects were asked to identify the excerpt's saxophonist given the choices of John Coltrane, Dexter Gordon, Sonny Rollins, or Wayne Shorter.

Results. Subjects were in general very good at identifying the performer playing in the excerpt. The average score was 11.9 out of 16 correct responses, where a chance level of performance would expect only 4 correct responses. Two of the excerpts were correctly identified by all subjects.

Conclusions. The results show that spectrographic cues can by themselves reveal a performer's identity. Such findings lead to important implications for jazz education, music analysis, and computational models of jazz improvisation, since these areas tend to emphasize the calligraphic aspects of jazz.

1. INTRODUCTION

How is it that listeners are able to recognize or distinguish one jazz performer from another? This article begins to formulate an answer to that question. An experiment is presented to support the hypothesis that well known jazz saxophonists can be recognized by their sound alone (i.e., without significant rhythmic or melodic information). As an illustration, consider the following procedure: Choose a solo by Coleman Hawkins, then (1) Remove the vibrato and the pitch bends, (2) Take every sixth note and transpose it randomly by ± 12 semitones, (3) Remove the 5th, 7th and 8th harmonics of every note, (4) Take every fifth note and transpose it randomly by between ± 3 semitones, (5) Take every seventh note and replace it with a rest, (6) Displace the onset of every other phrase by ± 150 milliseconds. At what point did Coleman Hawkins stop sounding like Coleman Hawkins? Does reordering the steps affect the outcome?

Jazz musicians have always acknowledged the importance of developing a unique stylistic voice as a way of transcending from imitation and assimilation into innovation (Berliner, 1994, p. 273). By dissecting the musical signal into its components in a controlled experiment setting, we can begin to assess which musical traits are most helpful to a listener attempting identification.

2. THE PERFORMER PROFILE

Two aspects of performance may be defined, what might be called *calligraphic* and *spectrographic*. This distinction will aid us in isolating the features that contribute to recognition. By *calligraphic*, I mean those aspects of the signal which can be represented via standard musical notation, namely, pitch, rhythm, and contour. "Zoomed-in" descriptions of these domains constitute the *spectrographic* category, which includes timbre (tone quality), microtemporal information (such as swinging and playing behind the beat), and expressive nuances (such as vibrato and pitch bending). Note that each step in the recipe of the opening paragraph altered a different feature of the signal, where step 1 = expression (S), 2 = contour (C), 3 = timbre (S), 4 = pitch (C), 5 = rhythm (C), 6 = microtiming (S).

2.1. Calligraphic Features

The number of analytical jazz articles that concentrate on calligraphic features far exceeds its spectrographic counterpart. This is probably a consequence of two factors: the influence of a pitch-oriented analytical tradition which has been useful in the context of Western European music, and (until recently) the technical difficulty involved in extracting and characterizing spectrographic features. The emphasis on calligraphic rather than spectrographic oriented discussion is especially rampant in jazz pedagogy. College-level courses in jazz theory typically devote the majority of their syllabus to scale types, chord progressions, and the relationship between the two. When jazz students refer to (say) a "bebop lick," they are generally alluding to a set of calligraphic features. The demand for calligraphic-type knowledge has been met by publishers in the form of innumerable solo transcriptions of virtually every major jazz musician. The opening sentence of Jerry Coker's widely read book, *Improvising Jazz* (1964), helps to underscore my point: "Five factors are chiefly responsible for the outcome of the jazz player's improvisation: intuition, intellect, emotion, sense of pitch, and habit." By treating pitch at a par with such colossal subjects as intellect and emotion (and by excluding sense of time, communication skills, etc.), Coker reflects and perpetuates the belief, however unfounded, that in jazz music calligraphic thinking reigns supreme.

ISBN 3-931852-67-9
ISSN 1617-6847

2.2. Spectrographic Features

Nonetheless, it would be misleading to maintain that writers have altogether underestimated the role of spectrographic features as determinants of a jazz performer's musical personality. When Duke Ellington (1980, p. 119) writes in his eulogy for alto saxophonist Johnny Hodges that his "tone [was] so beautiful it sometimes brought tears to the eyes," he encapsulates a general predisposition to highlight a performer's sound as the most characteristic feature of his or her playing. Similarly, the Grove Dictionary of Jazz attributes Hodges' playing authority to "the majesty of his *sound*, his endless vocabulary of *expressive ornaments*, and the *soulfulness* of his melodic ideas" – all spectrographic elements (italics added). In Gunther Schuller's (1968) list of qualities that set Louis Armstrong apart from the other improvisers of his age, calligraphic features are outnumbered by spectrographic features three-to-one. They are: Armstrong's (1) superior choice of notes, (2) quality of tone (i.e., timbre), (3) sense of swing (i.e., microtiming), and (4) repertory of vibratos and shakes (i.e., expression). Even though expressive nuances constitute an essential ingredient of a jazz performer's personal sound, scholars have so far avoided this area of research. Likewise, no scientific work has been done on the timbre profiles of jazz musicians, even though there exists an extensive literature on timbre and its perceptual attributes. In jazz circles, metaphors such as rough, raspy, sweet, bright, warm, full, fat, and dry are often used to characterize a particular player's sound. Schuller (1989, p. 859) scrapes the tip of the iceberg when he links quality of timbre to envelope characteristics. His analyses of the "envelope traces" of Louis Armstrong and bassist Ray Brown allow him to conclude that notes with "an envelope that starts gradually have a softer, springier, pure-sounding start. Notes that display an irregular envelope outline are likely to sound raspy or rough; where the envelope is smooth, the tones generally have a correspondingly smooth, uniform sound."

Of the three spectrographic categories discussed here, microtiming has drawn the most attention from jazz scholars. Using arrows and grids respectively, Haywood (1993) and Stewart (1982) developed notation techniques to represent the subtle temporal manipulations characteristic of jazz musicians. While graphically viewer-friendly, these strategies are limited in that they fail to provide accurate temporal data. Other studies have made use of spectrogram analysis to explore the role of microtemporal features in jazz, especially properties of long-short eighth-note patterns and the timing discrepancies within jazz ensembles (Ashley, 2002; Collier & Collier, 2002; Ellis, 1991; Friberg & Sundström, 2002; Prögler, 1995). Along the lines of this study, Benadon (2003) has shown that different jazz musicians employ contrasting approaches to beat subdivision, although it is still unclear how such microrhythmic profiles may influence the recognition task.

2.3. Describing Performers

When a listener attempts to characterize a performer, the resulting descriptions are typically referential, trait-oriented, or a combination of the two. Table 1 lists sample descriptions chronicled in Downbeat magazine's Blindfold Test, "a listening test that challenges the featured artist to identify the musicians who performed on selected recordings" (Downbeat). Such responses are informative but too broad to prompt definitive conclusions.

Response	Listener	Type
This is not Newk [Sonny Rollins], but there's some Newkisms in there.	James Carter	R
Reminded me of the way McCoy Tyner used to play.	Wayne Shorter	R
He has this amazingly unique, speedy vibrato.	Dave Liebman	T
His sound, the way he controls the time-feel.	Joe Lovano	T
The warm, arresting and caressing sound of Stan Getz.	Joe Henderson	C
I thought it was Art Tatum . . . I was waiting for that thunderous thing . . .	John Coltrane	C
There's a guttural thing in there that Ben [Webster] has.	James Carter	C

Table 1: Referential (R), Trait-Oriented (T), and Combination (C) responses in Downbeat Magazine's Blindfold Test.

3. THE RECOGNITION TASK

Musical analysis allows us to discover characteristic features of jazz musicians. What role do these features play during recognition? This is a difficult problem to tackle, primarily because our informal listening experience suggests that players are identifiably different for different reasons, so the extent to which each feature serves to provide clues may vary from player to player. Similarly, it is likely that different listeners employ different strategies to identify a given performer. For example, a listener who is more "sensitive" to microtiming manipulation might rely on this knowledge to guide the identification task, whereas a second listener might resort to his/her expertise with timbre profiles to identify the same listener. While these questions must await later studies for proper elucidation, we can take some initial investigative steps. The experiment presented below restricts the amount of information available to the listener, thus seeking to isolate the components that contribute to a positive identification.

3.1. Experiment

Seven volunteers participated in the study; the group was composed of university students and professors. All subjects considered themselves jazz connoisseurs, and were either current or former jazz performers on one of the following instruments: saxophone, piano, trumpet, and drums. The stimuli consisted of sixteen 2 to 3 second-long audio samples taken from jazz recordings of the 1950's and 60's (Table 2). Samples had the following characteristics. There was a brief (<100 ms) fade-in and fade-out; there was a tenor saxophonist playing in the foreground; the saxophonist was either John Coltrane, Dexter

ISBN 3-931852-67-9
ISSN 1617-6847

Gordon, Sonny Rollins, or Wayne Shorter; the saxophonist was heard playing between 2 and 5 notes; there were three instruments in the background: piano, upright bass, drums; meter was 4/4 and tempo was no slower than "medium" jazz tempo (i.e., no ballads); no two samples were taken from the same album (to minimize album-familiarity bias); tunes with famous accompaniment arrangements were avoided; samples were taken from part of a solo as opposed to the "head" (melody); samples were converted from stereo to mono to reduce left/right channel bias. Four different random sequences of the stimuli were prepared. Transposed (not "in C") transcriptions of the excerpts are shown in [SAXTRANSCRIPTIONS.GIF]. (The tenor saxophone is in Bb, sounding a major ninth lower than written.) Given the choices of John Coltrane, Dexter Gordon, Sonny Rollins, or Wayne Shorter, subjects were asked to identify the excerpt's saxophonist and circle his name on the experiment sheet. No information was given regarding the origin of the excerpts. Subjects were allowed to hear each excerpt up to five times. The session lasted between 15 and 20 minutes.

Player	Album / Track	Year
Coltrane	Crescent / Bessie's Blues	1964
Coltrane	Giant Steps / Cousin Mary	1959
Coltrane	Lush Life / I Love You	1957
Coltrane	Relaxin' With Miles / If I Were a Bell	1956
Gordon	Go! / Cheese Cake	1962
Gordon	One Flight Up / Kong Neptune	1964
Gordon	Our Man in Paris / Our Love is Here to Stay	1963
Gordon	A Swinging Affair / McSplivens	1962
Rollins	More Study in Brown / I'll Remember April	1956
Rollins	Plus 4 / Pent-Up House	1956
Rollins	Saxophone Colossus / Strode Rode	1956
Rollins	Sonny Side Up / The Eternal Triangle	1957
Shorter	Juju / Twelve More Bars to Go	1964
Shorter	A Night in Tunisia / So Tired	1960
Shorter	Speak No Evil / Witch Hunt	1964
Shorter	The Soothsayer / The Big Push	1965

Table 2: Origin of excerpts, each of which lasts between 2 and 3 seconds. See [SAXTRANSCRIPTIONS.GIF].

3.2. Results

Subjects were on average very good at identifying the performer playing in the excerpt. In descending order, the subjects' scores (out of 16) were 15, 15, 15, 14, 10, 8, 6. A chance level of performance would expect 4 out of 16 correct responses. The results suggest that spectrographic clues as brief as 2 to 3 seconds are sufficient for a positive identification. Figure 1 shows the total number of correct responses for each excerpt and saxophonist. Coltrane and Shorter received the highest scores, which is not surprising considering their popularity among jazz fans (some subjects confessed mid-experiment that they were not as familiar with Gordon or Rollins). Also, two excerpts by Shorter (S1 and S4) were correctly identified by all subjects. The excerpts containing the fewest notes (G4 and S3) received the lowest score. This suggests that listeners will squeeze as much information as possible – timbral or otherwise – out of each note during recognition. It also reminds us of the fact that, though considerably suppressed, calligraphic information has not been entirely removed. While none of the excerpts can hardly be considered as predominantly rhythmic or melodic, the contribution of calligraphic features during recognition may not have been entirely negligible. For example, the abundance of quarter notes in the Rollins excerpts differs noticeably from the more varied rhythmic configuration of the other performers.

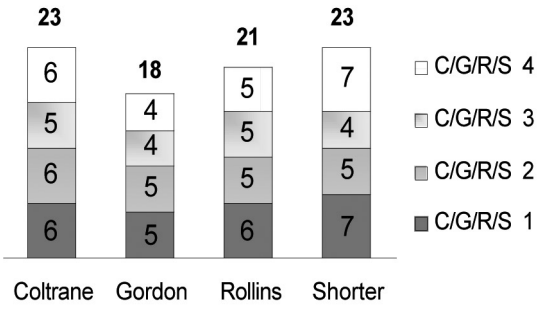

Figure 1: Number of correct responses per performer and excerpt.

The subjects' confusion matrices are shown in Figure 2. Seven out of eight Rollins errors consisted of selecting Gordon instead, but the reverse was not true: Gordon's confusion matrix is diffused (with a majority cluster around Coltrane). Also, none of the subjects mistook Rollins for Coltrane, nor Gordon for Shorter. Aside from these observations, no significant errors patterns emerged.

	DG	*SR*	*WS*		*JC*	*SR*	*WS*
C1			•	G1	•	•	
C2			•	G2	•	•	
C3		•	•	G3		•	••
C4	•			G4	•••		
	JC	*DG*	*WS*		*JC*	*DG*	*SR*
R1		•		S1			
R2		•••		S2			••
R3		••		S3	••		•
R4		•	•	S4			

Figure 2: Confusion matrices for Coltrane, Gordon, Rollins, and Shorter excerpts.

3.3. Remarks

We should be cautious to draw definitive conclusions from the subjects' incorrect responses. Recall that the sole criterion for selecting the participants was that they should consider themselves jazz connoisseurs. Clearly, this is a highly subjective way to determine a listener's degree of exposure to jazz – especially to tenor saxophonists of a particular era. The wide range of correct responses (6 to 15) across subjects implies that some of them

ISBN 3-931852-67-9
ISSN 1617-6847

had formed incomplete or "deficient" performer prototypes, thus diminishing efficiency during cue abstraction. Additionally, we should bear in mind that the difficulty of the subjects' task would increase considerably as the number of possible choices is increased. In keeping the performer pool reasonably small, not only did I facilitate lucky guesses, but I also allowed subjects to rely on processes of deduction and elimination. For example, if a subject is skilled at identifying Coltrane only, then an excerpt that clearly does not sound like Coltrane stands a 33% chance of being correctly identified – a considerable jump from 25%. No doubt the same thought process takes place in a "real life" recognition situation (i.e., given a very large number of possible answers), but the percentage increase for each step of elimination is minimal.

4. A FOLLOW-UP EXPERIMENT

Having shown that spectrographic features are powerful indicators of musical personality, it seems logical to conduct a follow-up study which assesses the role of calligraphic features during recognition. In a forthcoming experiment, subjects will listen to sixteen MIDI-piano transcriptions of phrases by the same saxophonists in the first experiment. These phrases are deemed by the author to be typical of each player (i.e., "licks"), and are not restricted as to the number of notes they contain. The stimuli should contain no dynamic changes, articulation, pitch bends, or time-feel information (i.e., they are quantized to the sixteenth-note); the MIDI-piano is the only instrument heard. Comparing the results of both experiments should shed some light on the interaction between spectrographic and calligraphic modes of recognition.

5. CONCLUSIONS

The experiment findings show that experienced listeners are able to identify different jazz saxophonists without the aid of substantial calligraphic information. This suggests that spectrographic cues can by themselves reveal a performer's identity, although it is not clear yet which of these cues – if any – carries the most weight during the recognition task. The implications raised by these results are especially relevant to the field of jazz pedagogy, which is disproportionately biased towards calligraphic-related topics.

6. REFERENCES

1. Ashley, R. (2002). Do[n't] change a hair for me: the art of jazz rubatos. *Music Perception, 19(3),* 311 – 332.

2. Benadon, F. (2003). The expressive role of beat subdivision in jazz. *Conference Proceedings of the Society of Music Perception and Cognition*, Las Vegas.

3. Berliner, P. F. (1994). *Thinking in jazz: The infinite art of improvisation*. University of Chicago Press: Chicago.

4. Coker, J. (1964). *Improvising jazz*. Simon & Schuster: New York.

5. Collier, G. L., & Collier, J. L. (2002). A study of timing in two Louis Armstrong solos. *Music Perception, 19(3),* 463 – 483.

6. Ellington, D. (1980). *Music is my mistress.* Da Capo Press: Garden City, NY.

7. Friberg, A., & Sundström, A. (2002). Swing ratios and ensemble timing in jazz performance: evidence for a common rhythmic pattern. *Music Perception,19(3),* 333 – 349.

8. Prögler, J. A. (1995). Searching for swing: participatory discrepancies in the jazz rhythm section. *Ethnomusicology, 39*, 21 – 54.

9. Schuller, G. (1968). *Early jazz: its roots and musical development*. Oxford University Press: New York.

10. Schuller, G. (1989). *The swing era: The development of jazz 1930 – 1945.* Oxford University Press: New York.

ISBN 3-931852-67-9
ISSN 1617-6847