

Commentary on Frieler, Pfeleiderer, Abeßer, and Zaddach's "Telling a Story"

YURI BROZE[1]

ABSTRACT: Theoretical and methodological perspectives are offered. Perceived personal involvement, variety of improvisational tools, and dramaturgy might be attributed to multimodal signals and cues of emotion, implicit learning of motor routines and musical tendency, and deliberate planning on the part of the musician. I give a personal perspective on my experience as an improvising musician, and suggest a sketch of how I imagine improvisation often works. Dramaturgical models meant for iteratively constructed artworks such as plays are likely to be deficient for improvisational artworks in general. Finally, the authors' methodological choices are considered. Both exploratory research and model selection are also driven by implicit hypotheses, so care must be taken to minimize false discovery.

Submitted 2015 December 24; accepted 2016 February 8.

KEYWORDS: *jazz analysis, improvisation, commentary, modeling*

IN a way, it is comforting that the primary finding in "Telling a Story" is simply that jazz solos tend toward increasing intensity. Imagine the horrifying alternative scenario: that the scholarly community somehow managed to perfectly describe every aspect of jazz improvisation. Such a discovery, tantamount to perfecting musical artificial intelligence, would be an academic windfall but an artistic crisis—to learn that this cherished form of expressive musicianship were completely deterministic might just doom the art form.

But this, of course, is not the intent. Instead, we seek to establish vocabulary and techniques that can be re-used in music creation, analysis, and yet further research. Every step forward we make in describing our art invites artists to think differently about their medium and how they approach it. If things go particularly well, we may also uncover knowledge about the human potential to create, to communicate, and to understand. In this capacity, "Telling a Story" is a fertile intuition pump and well-executed investigation into jazz solo structure. The authors first attack the problem using basic features of the musical surface (loudness and pitch), and subsequently with two different types of "preprocessing": algorithmic feature extraction at the level of musical phrase, and qualitative analysis of improvisational ideas.

Rather than to test explicit hypotheses, Frieler et al.'s study was intended to explore analytic techniques, to inspire ideas, and to suggest conjectures. I will offer some pursuant comments.

ON STORYTELLING AND DRAMATURGY IN JAZZ

Frieler and his colleagues believe that "good storytelling" in jazz is more rooted in aesthetic judgment (what people say) rather than in the musical artifact itself (what people play). Inspired by Berliner (1994), the authors outline three components of the "good storyteller aesthetic" as a launching point for their work. These are (a) personal involvement, (b) the variety and balance of improvisational tools employed, and (c) an overall dramaturgy.

If "musical storytelling" is indeed an aesthetic judgment, then we can regard "musical dramaturgy" as the craft of producing such judgments. Let us therefore reimagine these three components of perceived storytelling skill from the perspective of the performer's actions themselves. In essence, we're dealing with

(a) perceptual signals and cues of the performer's emotion, (b) aspects of musical syntax, rhythm, gesture, and tonality, and (c) overall plan for the solo.

Although this mapping is not perfect, it does suggest possible connections between each of the authors' three studies and a concrete causal agent for what is observable. The first study of basic surface features such as pitch height and loudness corresponds closely to vocal expression of emotion, and therefore we have a candidate explanans for perceived personal involvement. Likewise, their second phrase-wise analysis, which includes a number of mathematically derived structural and syntactic features, likely captures the residue of the performer's rehearsed motor routines and implicitly-learned expectations of musical structure. Last, leveraging as it does "improvisational ideas" of Schütz's (2015) midlevel analysis—*itself a method with reported external validity*—the third study seems ontologically seated somewhere in the conscious thoughts of the musician, at least to the extent that these ideas are planned several phrases in advance.

Placing empirical studies on solid generative footing may also help avoid the somewhat brutish equation of "tension curves" with "dramaturgy". Instead, dramaturgy could be imagined as a higher-level constellation of the availability of rehearsed motor routines, a repertoire of improvisational ideas, and any pre-improvisation planning of emotional or structural elements. These plans are physically executed with more or less fidelity, chance elements may enter the performance, and tension curves become things perceived rather than designed.

Against this backdrop, one can then conjecture that some improvising musicians sometimes choose to architect a solo's tension curve ahead of time, at some time before beginning the improvisation itself. This invites such questions as:

- How often do jazz musicians tend to pre-plan their choruses? Do they think in terms of tension curves?
- How often do apparently concave tension curves occur during planned versus unplanned solos? Is this possibly an emergent property of other concerns or causes? What impact does this have on performances?
- Do musicians who adopt such strategies tend to be judged as telling better stories? As more skilled?
- If a single musician deliberately changes their preplanning strategy, does this alter their perceived storytelling ability?

The degree to which premeditation and deliberate planning plays a role in the "dramaturgical" aspects of jazz soloing remains uncertain, but it seems likely that it would differ from musician to musician, performance to performance. The most effective techniques for following such lines of research would likely include surveys, interviews, and actual experimentation, rather than post-hoc analysis of existing improvisational artifacts.

ON IMPROVISATIONAL CONSTRUCTION

Improvisation's relationship to composition is somewhat akin to blazing a trail using a machete and compass rather than constructing an autoroute using backhoes and GPS. Like footpaths, improvisational solos are created primarily on-the-fly; compositions and freeways typically involve out-of-time iteration, layering, and refinement. Moreover, improvisational ideas often serve as kernels for fully developed compositions, as footpaths for paved roads. For their differences, it is doubtful at first glance that there would be any deep homology between dramaturgical models meant for the iterative process of playwrights and one apt for improvising jazz musicians. Nonetheless, Frieler and his colleagues point out the underlying overlap: both forms of performance traditionally rely on maintaining audience attention and interest.

If I were asked how I construct jazz piano improvisations, I would say that they are an amalgam of prior planning, melodic invention (with subsequent development and manipulation), true or pseudo-counterpoint, melodic and stylistic quotes, riffs, licks, hacks, and outright mistakes.[2] Anecdotally, these improvisations result from cognitive and motor plans, along with their respective malfunctions. Schütz (2015) has previously commented that many aspects of jazz improvisation ought to be understood in the context of the psychological present, and my intuition is consistent with this.

Despite having only absence of evidence rather than the converse from Frieler et al., it would certainly seem that the dramatic arc's influence over their studied population of jazz solos has quite a small effect size indeed. Considering that the sole dramaturgic similarity between jazz improvisation and theatre would seem to be the presence of an audience, and that jazz solos and plays are constructed under deeply different psychological states and epochs, we should remain open and even enthusiastic about constructing a separate theoretical model *de novo*.

ON EXPLORATORY RESEARCH AND MODELING

The distinction between exploratory and hypothesis-driven research is cultural, not essential; it is more a matter of rhetorical strategy than epistemology. Hypothesis-driven work embraces trepidation, guarding against overreach and false positives. Such a rhetoric is most useful for working out various incremental or fiddly bits after a theory has been established. In contrast, exploratory work is best suited for wide-open fields, where fruit hangs low and curious flashlights are newly shining. Yet in practice, exploratory research is still typically conjecture-driven, even if the conjectures are as simple as "perhaps something interesting is here," or even "wouldn't it be neat if we studied A using B?" Such exploratory work is conducted using small, iterative, informal hypotheses: hunches quickly acted upon, and subsequently pursued or discarded. The resulting report usually makes no definitive claims, but the trail is blazed, to recall the above metaphor, using similar statistical and inferential machinery.

Given that both forms of research are ultimately driven by a similar inductive process, it is understandable that the authors simultaneously claim their work as exploratory while also stating a soft hypothesis: that "[j]azz solos possess distinct curves of musical parameters that... might follow certain trends of dramatic development as, for example, an arched or concave curve." In this light, there are a few aspects of methodology worth reflecting on.

Choosing the form of a mathematical model in any empirical research paradigm is operationally equivalent to stating a hypothesis something like, "if a model with the functional form F is parameterized by variables B of cardinality P , and these variables B are computed from the data using technique L , then this model will be better (using metric Z) than all other models." Though one could quibble with this formulation, the point is that model selection itself requires the same methodological safeguards against false discovery as the rest of the work. If the model must be chosen using data, a data set independent of that used to fit parameters should be used. Although the authors are reassuringly sensitive to the problems of multiple testing, they still are vague about how, exactly, they experimented with other parameterizations and dynamic model selection. It would have been nice if this process took place using a separate data set entirely, in order to reduce the false discovery rate.

On the other hand, if model choice is driven by a priori knowledge alone, then the responsible methodologist will choose a form that is most likely to disprove their hypothesis. Testing hypotheses of arch-shaped patterns in any musical parameter is methodologically perilous, if only because "what goes up must come down." The null hypothesis of statistical regression to the mean must always be ruled out. In the worst case that the arch may occur in any arbitrary musical parameter, it becomes virtually tautological: surely all pieces would have some musical parameter that rises and then falls, especially when these parameters can be freely chosen.

At first glance, it is also somewhat troubling that the quadratic model was chosen, rather than one that could also disprove the arch hypothesis, say, by fitting additional complexity. Because arches are specifically sought, one would think that choosing an unconstrained quadratic model flirts with begging the

question. However, this concern falls away when looking at the time series themselves: it quickly becomes clear that the quadratic form is insufficient to capture whatever caused these variations in loudness and pitch, as Frieler and colleagues freely admit.

Instead, visual inspection of the pitch and loudness plots reveals what is potentially a rich hierarchical microstructure—and is also potentially mere noise, if the reader will excuse the pun. More work will be needed to unravel statistically detectable commonalities across solos and soloists. In so doing, future fitting of functional forms around melodic time series should be tilted toward musical, psychological, or physiological interpretability. Can a global quadratic be heard? I'm not so sure. Do jazz musicians regularly set out to create one when improvising? If they read *Telling a Story*, perhaps they will.

ON MULTIPLE TESTING AND FALSE DISCOVERY

Perhaps the distinguishing characteristic of Bayesian statistics is its reassignment of probability from the realm of mathematical limit conditions to the realm of subjective belief—that of the investigator. Whereas frequentist methods typically stop at producing an actionable statistic from the data (such as Type I error probability or a confidence interval), Bayesian techniques subsume the investigator's decision-making as well. It would be useful for all empirical researchers who use Bayesian techniques for hypothesis testing to apply them in awareness of this critical distinction.

Frieler and his colleagues claim to use Bayes Factors rather than the Bonferroni correction because they are interested in "general trends" rather than exact significance levels. This is fair enough, but the claim is also very much at odds with their choice to feature Bob Berg's solo for a prominent in-depth analysis. While musically interesting, there's no assurance that this solo is actually particularly different from the others.

Interpreting the Bayes Factor to learn what cannot "arise just by chance" as the authors do seems problematic, especially after having already admitted that these phrase-level parameters are correlated by mathematical construction. In the presence of extreme collinearity, the probabilistic implosion that usually accompanies independent trials simply does not occur. If the authors' binomially-computed expected fit numbers in Table 4 do not take into account these relationships, the Bayes Factors are likely inflated. Instead of using the theoretical binomial, a computerized permutation technique would be a more appropriate control.

CONCLUSION

Methodological cavils aside, Frieler and his coauthors have provided a very well-executed exploratory study of a corpus of jazz solos. In addition to providing a strong springboard for conjecture, they have demonstrated the utility of Schütz's midlevel analysis units when studying overall tendencies in jazz solo structure. This is particularly notable because of the technique's demonstrated external validity and potential usefulness in informing musical practice.

Unfortunately, much of this dataset's utility in testing theories of musical narrativity has now been spent on exploratory work. Future hypothesis-driven research will have to rely on independent sources of experimental or observational data.

NOTES

[1] Correspondence can be addressed to: yuri.broze+emr@gmail.com

[2] Produced as they are by a mediocre pianist, my solos usually gravitate toward the tail end of the list.

REFERENCES

Berliner, P. F. (1994). *Thinking in Jazz. The Infinite Art of Improvisation*. Chicago: University of Chicago Press. <http://dx.doi.org/10.7208/chicago/9780226044521.001.0001>

Schütz, M. (2015). *Improvisation im Jazz. Eine empirische Untersuchung bei Jazzpianisten auf der Basis der Ideenflussanalyse* (in German). Studien zur Musikwissenschaft, Vol. 34. Hamburg: Verlag.