The difficulty of discerning between composed and improvised music

ANDREAS C. LEHMANN*
REINHARD KOPIEZ**
*University of Music, Würzburg, Germany
**University of Music, Media and Drama, Hanover, Germany

● **ABSTRACT**

The generative processes (types) of composition and improvisation are often claimed to differ from each other with respect to complexity, spontaneity, and listener expectation. While this conceptual difference might be tenable, it is unclear whether it has an empirical perceptual basis. In a listening experiment with three contrasting pieces – classical (Mozart), romantic (Scriabin), and free jazz (Rissi) – we compared composed and non-composed musical stimuli. 102 experts rated six one-minute excerpts using the scales "spontaneous/improvised" and "coherent/rehearsed". The ratings were of acceptable, though moderate, reliability, and the listeners reacted markedly differently to the three styles. However, the effect of generative type was either non-significant (Mozart, Scriabin) or caused by performative aspects for the free jazz piece (Rissi). This lack of clear evidence for perceiving the two generative types is in line with other research showing divergence between theoretical concepts and their perceptual salience.

Keywords : composition - improvisation - creativity.

In an early review of the literature on generative processes, Sloboda (1985, pp. 102-150) suggested four sources of data to consider, namely information by composers about the compositional process, sketches, experimental data, and observations of improvised conduct. In this article, we have used experimental data to shed light on improvisation. However, we were interested in the listeners' viewpoints rather than the creators' perspectives. This research is in line with the work of Irène Deliège who has been concerned with listeners' structural or aesthetic perceptions and experiences (*e.g.*, Deliège, 1987; 2006). In the first part of the paper, we discuss the differences between improvisation and composition and listeners' awareness of structural aspects of music. In the next section, we report the findings of an exploratory listening experiment that we have conducted. A pilot study to this experiment was conducted in 2002 in Liège at the ESCOM conference organized by Irène Deliège (*cf.*, Lehmann & Kopiez, 2002).

In the musicological literature on improvisation and composition (for an overview see Seidel, 1987), several aspects recur that differentiate between the two "generative types", as we call them. First, improvisation is often characterized as being highly time constrained – *i.e.* it unfolds online in real-time, while composition happens mostly offline. Second, improvisation tends to be more of a patterned or formulaic type of music with a relatively high degree of redundancy, while compositions can achieve any desired amount of complexity. Third, improvisations are rarely written down, while compositions almost always exist in notation before they are executed. And finally, improvisations are characterized by a lower degree of rigor while compositions show a high degree of musical coherence. However, these four characteristics (time constraint, complexity, notation, coherence) may be less helpful for our understanding of musical creativity than is generally assumed (*cf.* Nettl, 1998; see Lehmann, 2005, for a comprehensive review on generative processes).

A clear theoretical distinction between improvisation and composition is difficult because both activities fulfill a similar function, namely, the intentional creation of an aesthetically satisfying and intentionally structured artifact. But is it possible for the audience to discern one from the other perceptually? This question has, at least to our knowledge, not been empirically tested. However, anecdotes regarding this problem do exist: The 19th century German music critic Adolph Bernhard Marx was once visited by the famous piano virtuoso Friedrich Kalkbrenner, who was kind enough to improvise for about 15 minutes at the piano. Marx was highly impressed and laudatory about the performance. A few days later, however, his admiration turned into anger when he received a package with new compositions from a publisher, among them the alleged improvisation he had heard a few days before. The work had already been composed and printed, and the critic deceived by the boisterous pianist (Gerig, 1974, cit. in Lehmann & Ericsson, 1998, p. 76). The avant-garde musician Carla Bley stated in an interview with improviser Peter Niklas Wilson that she did not like to draw a clear dividing line between composition and improvisation because the goal was to make the transition as seamless as possible, so that audiences do not know where the composition ends and the improvisation begins (*cit. in* Wilson, 1999, p. 163). In some cases, the difficulty of discerning between composition and improvisation can lead to innovative aesthetic concepts: for example, the Russian composer Nikolai Kapustin uses the idiom of jazz improvisation to compose piano works that sound improvised but are "notated down to the last hemidemisemiquaver [sixty-fourth note]" (Distler, 2004, p. 3). Kapustin's aesthetic idea is to integrate control of the material with decorative elements and rhapsodic features that the best jazz pianists bring to their playing. Also, there is the interesting case of composer Giacinto Scelsi. He tape-recorded and subsequently edited his transcribed instrumental improvisations, thereby causing controversy among critics who even considered his method compositional fraud (Fox &

Osmond-Smith, 2009). Also, as Berliner (1994) reported, in the area of jazz performance, especially in larger ensembles such as swing bands, arrangements are frequently notated and rehearsed even in parts where improvisation would be expected by the audience. Sometimes, improvisations are rehearsed to the point of a mental collective composition that may or may not be notated. Reports by individual musicians suggest that they practice and memorize parts of their improvisations (*cf.* Sidran, 1995).

One therefore has to consider the problem phenomenologically from the performer's point of view: Does he or she intend to play the same composition or reproduce a certain rendition, or does he or she intend to improvise? Nettl and Russell (1998) also suggest that, to the uninitiated listener of a particular ethnic music, different-sounding renditions of a piece may still be intended to be the same piece; conversely, same-sounding renditions could indeed be intended as different. Let us not forget that even meticulously rehearsed performances contain some degree of change from one rendition to the next due to the practice of memorized expressive performance cues (Chaffin, Lemieux, & Chen, 2006). Given that even expert listeners have few ways of distinguishing between composition and improvisation, one could surmise that the cognitive processes leading up to them are quite similar.

The mystery of the underlying processes for composition and improvisation has been heightened by recent fMRI studies. For example, Limb and Braun (2008) found that neural substrates of spontaneous musical performance – *i.e.* improvisation – were different from those of performing previously memorized music. And since the activation patterns showed an extensive deactivation in those regions typically associated with self-monitoring and conscious volitional control (*e.g.*, dorsolateral prefrontal and lateral orbital regions) as well as a deactivation of the limbic system, improvisation accompanied by sensorimotor activity on the instrument is probably different from the carefully planned compositional activity away from the instrument. Those findings may substantiate phenomenological accounts by composers, such as Aaron Copland who maintained,

> The inspired moment may sometimes be described as a hallucinatory kind of state of mind: one half of the personality emotes and dictates while the other half listens and notates. (…) Another kind involves the personality as a whole, or rather, loses sight of it completely, in a spontaneous expression of emotional release. (…) The less divine afflatus that makes it possible for us to compose each day – to induce inspiration, as it were – is a species of creative inspiration in which the critical faculty is much more involved (1952, p. 42-43).

Although this claim would require further empirical validation, we will focus here on the problem of discerning one product from the other perceptually.

Rather than dividing music in improvised and composed instances, it may be more helpful to differentiate the amount of what is given to the performer – the so-

called "referent" (*e.g.*, lead sheet, graphical notation, themes). The more detailed the referent, the more it predetermines the final product, and the less improvisation can take place. Pressing (1984, p. 349) provided a useful illustrative ranking of musical styles with their respective proportion of improvisation. At the top, we find traditional Japanese music and Western classical music which allow for minimal improvisation. At the bottom ranks Free Jazz with its many degrees of freedom. Since the degree of improvisation varies, it may only be noticeable – if at all – by listeners who are experts in a given style. No study to date has investigated this question.

Western music history of the last few hundred years has brought about an interesting dissociation between composer and performer, and this specialization cloaks some of the similarities inherent in both generative processes. The composer who is not at the same time a performer is a recent historical development, whereas composers in former times were professional also performers/improvisers. Due to the legacy of the aesthetic discourse of the 19th century, spear-headed by Eduard Hanslick (*cf.* Neitzert, 1991), today's audiences are largely unaware of the relatedness of compositional and improvisational activities for composers, such as Beethoven, Chopin, Liszt, and Bach. For example, Franz Liszt was not only famous for his superior pianistic but also for his vast improvisational skills. One of his master-class students reported an incident in which Liszt continued the performance of a rehearsed piece with a seamless transition into an improvisation: "Friedheim played some of Chopin's Preludes [Op. 28]. When he came to the ninth piece, the maestro was deeply moved by the religious Largo and suddenly said: 'I want to play this!' When he arrived at the last bar, he looked away from the score and started to improvise by prolonging the piece (...) to its threefold extent" (Lachmund, 1970, p. 147, [transl. RK]).

While modern expert musicians in the Western classical tradition have perfected their art of performing precomposed music, they have typically neglected performance of music by sight or through improvisation (*cf.* Lehmann, 2006). Over the years, the figured bass notation of baroque music has been replaced by notated arrangements provided by music theorists and composers; the cadenzas of solo concerts have been written out. Most concert goers therefore simply expect that classical and romantic music is composed and performed on stage from the score while jazz music, unless we are dealing with large ensemble arrangements for bigband, is improvised on stage.[1]

THE LISTENERS' ABILITIES

Ear training is a central part of every music education program, partly because musicians attempt to develop an analytical ear. Many students struggle with the demands of this subject in their degree programs. Extreme cases of autistic savants show to what extraordinary extent our analytical senses can be honed and connected

---

1 The authors of this article are well aware that a change in tide is happening and many musicians, especially those adhering to early music performance practice, do improvise.

with psychomotor skills (*e. g.*, Ockelford, 2007). However, several researchers have been able to show that even skilled listeners are not always able to deliver the performance they think they can deliver. For example, in Cook's (1987) ground-breaking experiment, even music students could not reliably discriminate between the correct version of a piece with its recapitulation in the fundamental key and a manipulated version with the recapitulation in a distant key, when the time span between the beginning of the piece and the recapitulation was longer than one minute. These findings indicate that means of compositional organization, such as large-scale formal functions of tonality, "may not in [themselves] be perceptible" (Cook, 1987, p. 197; see also Discussion for further comments). Or, in a reliability study on instrument quality ratings by professional trumpet players, Bertsch and Waldherr (2005) found mixed results: Some rating categories showed high reliability (*e.g.*, overall instrument quality), while other categories were only of limited reliability (*e.g.*, intonation, personal preference, response, air consumption). Support for this relativistic view can also be found in a recent study of one of the authors. Kopiez and Platz (2009) showed that only 49% of their subjects (music students) noticed a clash of keys between the melody and accompaniment in a song, when their attention was not directed toward the dissonance. Yet, when their attention was focused on the relationship between melody and accompaniment, 78% of the subjects noticed a clash of keys. In another study, Ziv and Moran (2006) insinuated that one of two synthesized performances of a Chopin prelude was played by a human performer, the other by a computer. Listeners consistently rated the alleged human performance to be superior on all accounts. Finally, in an informal pilot study to this paper, we presented two pieces of music to an audience who had attended a talk at a conference (see Lehmann & Kopiez, 2002). We then asked listeners to judge whether the examples were improvised or composed. The counted votes were roughly equally distributed between the alternative answers. This led us to believe that even experts, in this case our esteemed colleagues, encounter difficulties in discerning between both generative types. Thus, listeners' attention, (situational) expectations, experiences, and biases influence what they hear or believe they hear.

<center>**METHOD**</center>

PARTICIPANTS AND PROCEDURE
Participants in this listening experiment consisted of 102 German college music students of typical undergraduate age and experience. They were enrolled in music psychology or systematic musicology courses as part of their regular degree programs. In groups, listeners were presented six musical excerpts, of which half were composed and the remaining half were non-composed or improvised in some way. The order of presentation of the stimuli was balanced between groups. The instructions referred

<center>**117**</center>

to six interpretations that the subjects were to judge on five-point unipolar rating scales by marking their agreement to the scales "coherent/orderly", "spontaneous/ improvised", "deliberated/rehearsed". Preference and familiarity for style and the particular piece were also collected. A subgroup of the listeners heard a repeated presentation of some of the pieces to assess reliability of the judgments.

STIMULI

Three composed pieces of music were carefully selected based on the availability of a close match that could be called non-composed or improvised (see Table 1).[2] The first was a short classical minuet for piano by Wolfgang A. Mozart (minuet no. 1 from piano sonata KV 282, performed by Alfred Brendel, 1992); the second was the sonata no. 9, Op. 68, "Black Mass", by Alexander Scriabin (recording by Vladimir Ashkenazy, 1989); the third was a free jazz piece for ensemble entitled "Papago" by Mathias Rissi, Guerrino Mazzola, and Heinz Geisser (1996). Each piece was cut to an excerpt of about one minute in length. As mentioned previously, there were non-composed matches to each piece. The minuet was part of a collection of musical dice games that famous composers like Mozart and Haydn wrote in the second half of the 18th century. With the help of the electronic version of Mozart's own musical dice game (KV294d/516f; Op. posthumus, 1793), a new minuet/waltz was generated by a random recombination of one-bar musical fragments into suitable alternative instances (Reuter, 2001). A match for the section from the sonata by Scriabin was taken from a recording of improvisations by the German pianist Guenter Philipp (1999) on the same sonata, who specialises in improvisations in the idiom of Scriabin's piano music. Finally, the free jazz improvisation "Papago" contains a stretch of composed music in the beginning, allowing for a direct contrast with an improvised section within the same piece (G. Mazzola, personal communication, March 2002).

Table 1.

Musical examples used in the listening experiment (excerpts of approx. 60 seconds).

	Generative type	
Style and composer	*Composed*	*Non-composed ("Improvised")*
Classical (Mozart)	Minuet KV 282 (original) (Brendel, 1992)	Dice game KV294d/516f; Op. posthumus (Reuter, 2001)
Romantic (Scriabin)	Sonata No. 9 (Op. 68) (Ashkenazy, 1989)	Improvisation on Scriabin motives (Philipp, 1999)
Free Jazz (Rissi *et al.*)	"Papago" (composed section)	"Papago" (improvised section)

2 Sound examples of the pieces used for the listening experiment can be obtained from the website http://musicweb.hmtm-hannover.de/impcomp/

RESULTS

The analyses were carried out in two steps. First, we explored the relationship between the three relevant unipolar scales (coherent/orderly, spontaneous/improvised, deliberated/rehearsed) and then examined the influence of musical style and generative type on these dependent variables.

The initial exploratory data analysis revealed that for the dependent variables coherent/orderly, spontaneous/improvised, and deliberated/rehearsed, there was virtually no variability in responses for the Mozart examples, rendering parametric statistics for these two versions inappropriate. We therefore excluded these examples from the initial parametric analyses but will return to them later in the report of results.

Correlations between the three main scales yielded results of expected direction but somewhat larger than expected magnitude (see Table 2). Spontaneous/improvised was negatively correlated with coherent/orderly (r = -.59) and deliberated/rehearsed (r = -.68). Coherent/orderly and deliberated/rehearsed were positively associated (r = .60) Therefore, we averaged coherent/orderly and deliberated/rehearsed to compute a new and more stable measure named "coherent/rehearsed". This new variable also compensated somewhat for the moderate reliabilities reported below. Consequently, the resultant variable and spontaneous/improvised were highly negatively correlated (r = -.71). Our dependent variables in the following analyses were therefore the two variables spontaneous/improvised and coherent/rehearsed, which appear to capture opposing characteristics of our stimuli.

Table 2

Correlations between original variables (1, 2, 3) and newly combined variable (4)

	1. coherent/orderly	2. spontaneous/improvised
2. spontaneous/improvised	-.59 (408)	
3. deliberated/rehearsed	.60 (407)	-.68 (407)
4. coherent/rehearsed (averaged from 1 and 3)	--	-.71 (408)

Note. All correlations significant at p < .01 level; the correlations exclude data for the Mozart examples (see Text for details).

A subsample of 40 listeners received repeated exposure to one of the pieces (Scriabin composed). The retest-reliability was moderate with correlations of $r(40)$ = .52 (p < .01) for coherent/orderly, $r(39)$ = .42 (p < .01) for spontaneous/improvised, $r(39)$ = .40 (p < .02) for deliberated/rehearsed, $r(39)$ = .63 (p < .01) for preference, and $r(40)$ = .60 (p < .01) for familiarity with the style. One could argue that perceived coherence and spontaneity change with repeated exposure while preference for a piece and familiarity with the style might remain constant. The reliabilities are not high, but acceptable.

In the next repeated measures analyses using the general linear model of SPSS (version 17), we assessed the influence of the within-subject variables generative type (levels: composed, non-composed) and style (levels: Scriabin, Rissi), separately for our two dependent variables (coherent/rehearsed and spontaneous/improvised). As mentioned above, the Mozart pieces had to be excluded due to lack of variability on some scales. The hypothesis had been that the effect of style would be significant because the pieces were sufficiently different, with the chronological order (romantic, modern) offering a plausible sequence of values for judgments of coherence and spontaneity. The generative type was of specific interest because an effect here would show that listeners detected differences in structure with regard to our dependent variables. The directed hypothesis was that listeners would experience improvisations as more spontaneous/improvised while compositions could appear more coherent/ rehearsed.

The mean *coherent/rehearsed* scores for the Scriabin and Rissi pieces for the "generative type" factor differed significantly beyond the .01 level: $F(1,101) = 86.3$, $p < .001$, with a large effect size of partial eta-squared = .46 (see Tables 3 and 4). The mean ratings for the effect of the "piece" factor also differed significantly at the .01 level: $F(1,101) = 315.5$, $p < .001$, yielding a large effect with partial eta-squared of .76. The "piece by generative type" interaction was also reliable, $F(1,101) = 40.2$, $p < .001$, showing a partial eta-squared of .28 (see Figure 1).

Table 3.

Summary of repeated measures analysis regarding generative type and style for dependent variable "coherent/rehearsed".

Source	SS	df	MS	F	p	Effect size
Style	208.26	1	208.26	315.5	< .001	.76
Generative type	31.02	1	31.02	86.3	< .001	.46
Style * Generative type	22.35	1	22.35	40.2	< .001	.28
Within cells (error)	106.0	101	1.05			

Note. Effect size > .14 = large (*cf.* Sedlmeier & Renkewitz, 2008, p. 479).

Table 4.

Means and standard deviations for generative type condition as a function of style for dependent variable "coherent/rehearsed".

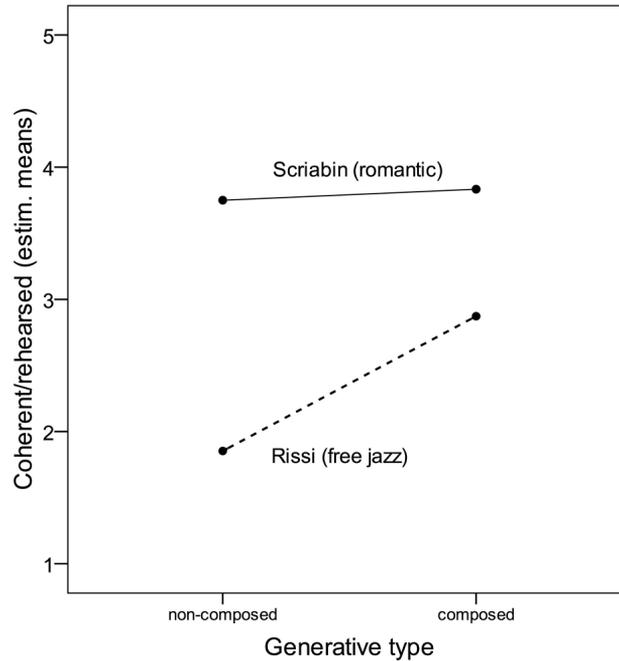| Style	Generative type *composed*		Generative type *non-composed*	
	M	SD	M	SD
Scriabin	3.8	.75	3.7	.76
Rissi	2.9	.95	1.8	.75

Note. N = 101

Fig. 1.

Means for interaction effect between style and generative type for the dependent variable coherent/rehearsed.

The mean *spontaneous/improvised* scores for the Scriabin and Rissi pieces for the "generative type" factor differed significantly beyond the .01 level: $F(1,101) = 21.5$, $p < .001$, with a large effect size of partial eta-squared = .17 (see Tables 5 and 6). Also, the mean ratings for the "style" effect differed significantly at the .01 level: $F(1,101) = 142.3$, $p < .0$ elding a large effect with partial eta-squared of .58. And the "piece by generative type" interaction was reliable, $F(1,101) = 21.1$, $p < .001$, showing a partial eta-squared of .17 (see Figure 2).

Table 5.

Summary of repeated measures analysis regarding generative type and style for dependent variable "spontaneous/improvised".

Source	SS	df	MS	F	p	Effect size
Style	225.02	1	225.02	142.3	< .001	.58
Generative type	18.55	1	18.55	21.5	< .001	.17
Style * Generative type	21.20	1	21.20	21.1	< .001	.17
Within cells (error)	213.8	101	2.1			

Note. Effect size > .14 = large (*cf.* Sedlmeier & Renkewitz, 2008, p. 479).

Table 6.

Means and standard deviations for generative type condition as a function of style for the dependent variable "spontaneous/improvised".

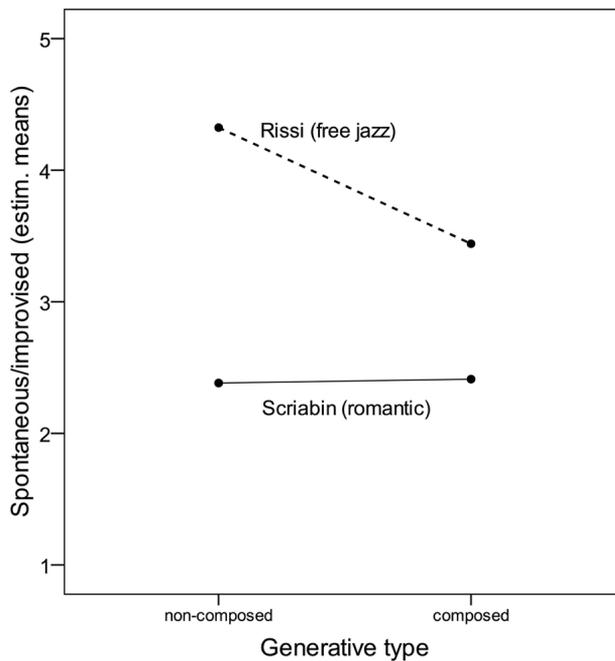| Style	Generative type *composed*		Generative type *non-composed*	
	M	SD	M	SD
Scriabin	2.4	1.09	2.38	1.20
Rissi	3.4	1.30	4.32	1.30

Note. N = 101

Fig. 2.

Means for interaction effect between style and generative type for the dependent variable spontaneous/improvised.

Using a paired sample t-test, we analyzed the differences between composed and non-composed versions for the spontaneous/improvised and the coherent/rehearsed scales. As Figures 1 and 2 suggest, the results were significant for Rissi (t[101] = 5.9, p < .01 for spontaneous/improvised and t[101] = 10.8, p < .01 for coherent/ rehearsed). The results for Scriabin were not reliable. Thus, the significant "generative type" and "generative type by style" effects are solely due to the differences in the Rissi piece. Finally, knowledge of the piece as a covariate did not lead to significant results, suggesting that knowing the piece did not matter.

For the Mozart piece we analyzed the effect of generative type, this time using non-parametric statistics. Two separate Mann-Whitney tests were conducted, one for each dependent variable. The mean ratings for *coherent/rehearsed* for the non-composed version (M = 4.70, SD = .47, 95% CI = 4.61/4.80) were not reliably higher than those for the composed version (M = 4.69, SD = .61, 95% CI = 4.56/4.80). Similarly, non-significant results were obtained for the scale *spontaneous/ improvised* for the non-composed (M = 1.20, SD =.53, 95% CI = 1.09/1.30) and for the composed version (M = 1.35, SD = .83, 95% CI = 1.18/1.51). Thus, the results of the Mozart comparison follow closely those found for the Scriabin pieces and would lead to a similar graphical representation (see Scriabin in Figure 1 and 2 for illustration).

In sum, our main dependent variables spontaneous/improvised and coherent/ rehearsed were highly negatively correlated. Our listeners distinguished readily between styles, but did not assign differential ratings to matched composed and non-composed versions of the classical (Mozart) and romantic (Scriabin) music pieces. The free jazz (Rissi) composed and non-composed excerpts were rated differentially, although they originated from the same piece.

Discussion

Our theoretical hypothesis was that the two generative processes (or types) leading to composition and improvisation are much more akin than commonly assumed. The large variability in observable behavior (*e.g.*, process traces from different composers) and phenomenology (*e.g.,* self-reports of composers or improvisers) with regard to the two generative types may prompt us to postulate differences between improvisation and composition that, in fact, have little perceptual validity. Adding to this, music history has taught us that classical music is composed while jazz is improvised, thus creating a confound of generative type and style. Finally, as concerns at least the live situation of jazz performance, the audience may participate in a conventionalized collaboration with the performers to act as if the music is being created at that moment and worthy of appreciation (Auslander, in press).

Given that musical styles such as classical, rock or folk are quickly identified in everyday life, (within a quarter second; *cf.* Gjerdingen & Perrott, 2008) but no distinction is made between improvised and composed music (*e.g.*, radio stations do not specifically announce whether a piece is improvised or not), we wanted to show under controlled conditions that expert listeners cannot easily discern one generative process from the other. Obviously, this is an epistemological difficulty because we were aiming for a null result.

In a study with three carefully selected pieces of composed music and their matched non-composed counterparts, listeners rated the interpretation of music excerpts (classical, romantic, and free jazz music) on the adjectives scales "spontaneous/

improvised" and "coherent/rehearsed". That the listeners would be sensitive to the different styles was expected and confirmed. More interesting was the fact that the distinction of generative types – that is, composed versus non-composed – for our classical (Mozart) and romantic (Scriabin) examples was not significant. Discernment appears difficult, if not impossible, when the non-composed stimulus is of high quality and played on a single instrument. Yet, listeners succeeded in the case of the free jazz excerpts (Rissi): The composed version was judged to be reliably less spontaneous/improvised and more coherent/rehearsed than the non-composed one. This is surprising because this was the piece with the least constraints and the strongest structural resemblance of the two types. What could be the reason?

Deliège talks of cues and the extraction of such cues by listeners. She also follows the ideas of Rosch in her writing on prototypes (Deliège, 2006). The question one could pose when relating our experiment to the work of Irène Deliège is, what cues are listeners using to discern improvisation from composition, and do they have separate categories for those two generative types? The answer we venture to give after our exploratory work is the following: Listeners do not have clear categories for improvisation and composition, but they can search for cues that may hint at the underlying generative processes. However, contrary to what musicologists or theorists might think in the aesthetic tradition of Western art music, the cues do not reside in the structural features of the work but rather emerge from its performative character. Improvisation suggests freedom which can be heard in particular in the rhythmic embodiment of the music. More constraints and "togetherness" of an ensemble may indicate composition, while a higher degree of entropy could signal improvisation. This cue is more easily extracted from performances with several instruments as was the case in the Rissi example: When played from notation, the ensemble will sound more together, whereas during the improvisation the rhythms are more freely executed.

Much future research could be done. Our tentative explanation requires further experimentation in which rehearsed pieces are presented with varying degrees of performative freedom. Given the limited scope of our choice of pieces and their length, a next step could also be the production of further stimuli by asking performers to play a piece and then improvise on it (see also anecdote about Liszt earlier). Performers, especially those in the early music arena, have internalized idiomatic features of a given style and can skillfully generate works in real-time that, to the uninitiated listener, are virtually indistinguishable from compositions.[3] Also, using longer excerpts may be useful to allow for more large-scale structures. In light of research showing fast detection of styles (Gjerdingen & Perrott, 2008) and emotional expression (within 600 ms, see Goydke, Altenmüller, Möller & Münte,

---

3 One of us (AL) was invited to give a talk at the Schola Cantorum Basiliensis and was privy to a reconstruction of a historical „Orgelprobe" (try-out for an organ post). The organ players executed the most marvellous improvisational feats!

2004), it is unclear how long the excerpts should be. For example, improvisatory classical music from India and the Middle East uses lengthy pieces with very gradual developments and interspersed compositions. Finally, it may be useful to include audio-visual material to broaden the scope toward ensemble performance, because improvisation in groups requires (facial and bodily) communication.

Our conclusions align with those of Cook (1987) concerning musical concepts and percepts. Even if the principles of compositional organization, here generative types, are not always perceptible, they are nonetheless important concepts: "(...) musicians sometimes forget the converse fact that what is good music theory may be bad psychology (...). In reminding musicians of this, the experimental psychologist can provide a valuable, if not always welcome, service" (Cook, 1987, p. 205).

AUTHOR NOTES

Address for correspondence:
Andreas C. Lehmann
School of Music, Hofstallstr.
6-8, 97070 Würzburg
Germany
Email: ac.lehmann@hfm-wuerzburg.de

- **References**

Ashkenazy, V. (1989). *Scriabin piano sonatas* [CD]. London: Decca.

Auslander, P. (in press). Jazz improvisation as a social arrangement. In N. Cook & R. Pettengill (Eds.), *Music as performance: New perspectives across the disciplines.* Ann Arbor: Michigan University Press.

Brendel, A. (1992). *Mozart: Sonatas KV533/494 & 282* [CD]. Hollywood: Philips Digital Classics.

Berliner, P. A. (1994). *Thinking in jazz.* Chicago: Chicago University Press.

Bertsch, M., & Waldherr, K. (2005). Reliability of trumpet players' judgments during instrument blind-tests. In *Proceedings of the Second Congress of Alps-Adria Acoustics Association and First Congress of Acoustical Society of Croatia* (pp. 207-216), Opatija, Croatia, 23-24 June.

Chaffin, R., Lemieux, A. F., & Chen, C. (2006). Spontaneity and creativity in highly practised performance. In I. Deliège & G. A. Wiggins (Eds.), *Musical creativity* (pp. 200-218). Hove, UK: Psychology Press.

Cook, N. (1987). The perception of large-scale tonal closure. *Music Perception, 5*(2), 197-206.

Copland, A. (1952). *Music and imagination.* Cambridge: Harvard University Press.

Deliège, I. (1987). Grouping conditions in listening to music: an approach to Lerdahl and Jackendoff's grouping preference rules. *Music Perception, 4*(4), 325-360.

Deliège, I. (2006). Creative support to elaborate a model of music listening. In I. Deliège & G. A. Wiggins (Eds.), *Musical creativity: Multidisciplinary research in theory and practice* (pp. 63-77). Hove: Psychology Press.

Distler, J. (2004). *Nikolai Kapustin piano music* [CD booklet]. London: Hyperion

Eisenberg, J., & Thompson, W. F. (2003). A matter of taste: evaluating improvised music. *Creativity Research Journal, 15*(2/3), 287-296.

Fox, C., & Osmond-Smith, D. (2009). "Scelsi, Giacinto." In *Grove Music Online.* Oxford Music Online, (accessed December 27, 2009).

Gjerdingen, R. O., & Perrott, D. (2008). Scanning the dial: The rapid recognition of music genres. *Journal of New Music Research, 37*(2), 93-100.

Goydke, K., Altenmüller, E., Möller, J., & Münte, T. (2004). Changes in emotional tone and instrumental timbre are reflected by the mismatch negativity. *Cognitive Brain Research, 21*(3), 251-359.

Kopiez, R., & Platz, F. (2009). The role of listening expertise, attention, and musical style in the perception of clash of keys. *Music Perception, 26*(4), 321-334.

Lachmund, C. V. (1970). *Mein Leben mit Franz Liszt. Aus dem Tagebuch eines Liszt-Schülers* [My life with Franz Liszt. From the diary of a Liszt student] Eschwege: Schröder.

Lehmann, A. C., & Ericsson, K. A. (1998). Historical developments of expert performance: public performance of music. In A. Steptoe (Ed.), *Genius and the mind* (pp. 67-94). Oxford: Oxford University Press.

Lehmann, A. C., & Kopiez, R. (2002). Revisiting composition and improvisation with a historical perspective. In M. Britta, & M. Mélen (Eds.), *Proceedings of the 10th Anniversary Conference of the European Society for the Cognitive Sciences of Music (ESCOM): Musical Creativity*, Liège, Belgium, 5-8 April [CD-ROM].

Lehmann, A. C. (2005). Komposition und Improvisation: Generative musikalische Performanz [Composition and improvisation: Generative musical performance]. In T. H. Stoffer & R. Oerter (Eds.), *Allgemeine Musikpsychologie* (pp. 913-954). Göttingen, Germany: Hogrefe.

Lehmann, A. C. (2006). Historical increases in expert music performance skills: optimizing instruments, playing techniques, and training. In E. Altenüller, M. Wiesendanger & J. Kesselring (Eds.), *Music, motor control, and the brain* (pp. 3-22). Oxford: Oxford University Press.

Limb, C. J., & Braun, A. R. (2008). Neural substrates of spontaneous musical performance: An fMRI study of jazz improvisation. *PLoS ONE, 3*(2), e1679. doi:10.1371/journal.pone.0001679.

Neitzert, L. (1991). Der Komponist, sein Werk und die subversive Kraft des Extempore [The composer, his work, and the subversive power of extemporization]. In W. Knauer (Ed.), *Jazz und Komposition* (pp. 37-53). Hofheim, Germany: Wolke.

Nettl, B. (1998). Introduction: An art neglected in scholarship. In B. Nettl & M. Russell (Eds.), *In the course of performance* (pp. 1-23). Chicago: Chicago University Press.

Nettl, B., & Russell, M. (Eds.) (1998). *In the course of performance.* Chicago: Chicago University Press.

Ockelford, A. (2007). A music module in working memory? Evidence from the performance of a prodigious musical savant. *Musicae Scientiae,* Special Edition on Performance, 5–36.

Philipp, G., & Pruggmayer-Philipp, U. (1999). *Klavierimprovisationen* [CD]. Berlin: Eterna.

Reuter, C. (2001). *Musikalische Würfelspiele von Mozart, Haydn und anderen großen Komponisten* [CD-ROM] [Musical dice games by Mozart, Haydn and other great composers]. Mainz: Schott.

Rissi, M., Mazzalo, G., & Geisser, H. (1996). *Fuego* [CD]. Root, Switzerland: Creative Works.

Sedlmeier, P., & Renkewitz, F. (2008). *Forschungsmethoden und Statistik in der Psychologie* [Research methods and statistics in psychology]. München: Pearson Studium.

Seidel, W. (1987). *Werk und Werkbegriff in der Musikgeschichte* [Works and the concept of works in music history]. Darmstadt: Wissenschaftliche Buchgesellschaft.

Sloboda, J. A. (1985). *The musical mind.* Oxford: Clarendon.

Sidran, B. (1995). *Talking jazz. An oral history.* New York: Da Capo.

Sternberg, R. (Ed.) (1999). *Handbook of creativity.* Cambridge: Cambridge University Press.

Wilson, P. N. (1999). *Hear and now.* Hofheim, Germany: Wolke.

Ziv, N., & Moran, O. (2006). Human versus computer: The effect of a statement concerning a musical performance's source on the evaluation of its quality and expressivity. *Empirical Studies of the Arts, 24*(2), 177-191.

• La dificultad de discernir entre música compuesta e improvisada

Se dice que los procesos generativos (tipos) de composición e improvisación difieren entre sí en términos de complejidad, espontaneidad, y la expectativa en el oyente. Si bien se podría argumentar que tal diferencia conceptual existe, no está claro si hay bases empíricas en el dominio de la percepción para sostener tales afirmaciones. En un experimento auditivo con tres piezas contrastantes - clásica (Mozart), romántica (Scriabin), y *free* jazz (Rissi) - comparamos estímulos musicales compuestos y no-compuestos. 102 expertos calificaron seis extractos de un minuto, usando las escalas "espontáneo / improvisado" (*spontaneous / improvised*) y "coherente / ensayado" (*coherente / rehearsed*). Las calificaciones fueron aceptables, aunque de moderada fiabilidad, y los oyentes reaccionaron notablemente diferente a los tres estilos. No obstante, el efecto del tipo generativo fue no significativo (Mozart, Scriabin) o causado por aspectos relacionados con la ejecución de la pieza de free jazz (Rissi). La falta de evidencia clara respecto a la percepción de los dos tipos generativos está en línea con otras investigaciones que demuestran la divergencia entre los conceptos teóricos y su relevancia perceptiva.

• La difficoltà di discernere tra musica composta ed improvvisata

Spesso si presume che i (tipi di) processi generativi nella composizione e nell'improvvisazione differiscano tra loro per complessità, spontaneitá ed aspettativa nell'ascoltatore. Se pure questa differenza concettuale è accettabile, risulta meno chiaro se essa abbia una base percettiva empirica. In un esperimento di ascolto con tre pezzi contrastanti – uno classico (Mozart), uno romantico (Scriabin) e uno di "free" jazz (Rissi) – abbiamo confrontato stimoli musicali composti e non-composti. 10 esperti hanno giudicato sei estratti di un minuto usando le scale "spontaneo/improvvisato" e "coerente/eseguito". I giudizi erano da accettabile fino a moderata affidabilità e gli ascoltatori reagivano in modo marcatamente diverso ai tre stili musicali. Comunque, l'effetto del tipo generativo o non era significativo (Mozart, Scriabin) oppure era causato dagli aspetti esecutivi del pezzo di "free" jazz (Rissi). Quest'assenza di una dimostrazione chiara della percezione di due tipi generativi è in linea con precedenti ricerche che provano la divergenza tra concetti teorici e la loro salienza percettiva.

• De la difficulté de discerner entre musique composée et musique improvisée

Les processus (ou types) génératifs associés à la composition et à l'improvisation sont souvent posés comme différents, aussi bien du point de vue de la complexité, de la spontanéité que des attentes des auditeurs. Bien que cette différence conceptuelle soit raisonnable, son fondement perceptuel n'a pas été mis en évidence de manière empirique. Dans une expérience d'écoute basée sur trois pièces contrastées – de styles respectivement classique (Mozart), romantique

(Scriabine) et *free jazz* (Rissi) – nous avons comparé des stimuli musicaux composés et non-composés. 102 experts ont annotés six extraits musicaux d'une minute le long des échelles « spontanée / improvisée » et « cohérent / répété (*rehearsed*) ». Les réactions des auditeurs offrent une consistance acceptable, bien qu'assez modérée, et les auditeurs ont réagi de manière sensiblement différente en fonction des trois styles. Toutefois, l'effet des types génératifs est soit non-significatif (Mozart, Scriabine), soit, dans le cas de la pièce de free jazz (Rissi), provoqué par des aspects du jeu musical. Ce manque d'évidence claire concernant la perception des deux types génératifs est en ligne avec d'autres recherches montrant la divergence entre les concepts théoriques et leur saillance perceptuelle.

• Die Schwierigkeit, zwischen komponierter und improvisierter Musik zu unterscheiden

Es wird häufig behauptet, dass sich die generativen Prozesse von Komposition und Improvisation in Hinblick auf Komplexität, Spontanität und Zuhörer-Erwartungen unterscheiden. Während sich die Unterschiede auf konzeptioneller Ebene als haltbar erweisen mögen, ist es unklar, ob derartige Unterschiede auf perzptueller Ebene empirisch erfassbar sind. In einem Hörexperiment verglichen wir mittels dreier kontrastierender Hörbeispiele – ein klassisches (Mozart), ein romantisches (Skrijabin) und ein Free-Jazz-Hörbeispiel (Rissi) – komponierte und nicht-komponierte Musik miteinander. Experten (N = 102) bewerteten einminütige Ausschnitte (N = 6) auf den Skalen „spontan/improvisiert" und „kohärent/einstudiert". Die Bewertungen wiesen eine annehmbare, wenn auch moderate Reliabilität zwischen den Urteilern auf. Die Zuhörer reagierten zwar deutlich unterschiedlich auf die drei Stile, jedoch war die Wirkung eines generativen Typs entweder statistisch unbedeutend (Mozart, Skrijabin) oder im Fall des Free-Jazz-Hörbeispiels (Rissi) mit Hilfe performativer Aspekte erklärbar. Dieser klare Mangel an Beweisen für die wahrnehmungsbasierte Unterscheidung zwischen zwei generativen Typen stimmt mit anderen Forschungsergebnissen überein, welche die Divergenz zwischen theoriegeleiteten Konzepten und deren perzeptueller Auffälligkeit belegen.