journal of interdisciplinary music studies
spring 2012, volume 6, issue 1, art. #12060101, pp. 01-21

Expressive performance cues in Gypsy swing guitar style: A case study and novel analytic approach

Marko Aho[1] and Tuomas Eerola[2]

[1] School of Social Sciences and Humanities, University of Tampere
[2] Finnish Centre of Excellence in Interdisciplinary Music Research, University of Jyväskylä

Background in ethnomusicology. In ethnomusicology the emotional expressiveness of musical performances still remains a minimally discussed topic. The expansion of local styles such as gypsy swing to global scales prompts one to seek cross-cultural bodily gestures behind the expressive conventions of musical styles; furthermore, such expressive conventions may be guided by the physical features of musical instruments.
Background in cognitive musicology. Considerable research in music cognition has been carried to identify the expressive performance cues. However, most studies concentrate on conveniently measurable properties of the instruments and on performance of composed works that have been recorded multiple times. Expression in improvisational music has not received much attention yet despite the broader range of possibilities it offers for the expressive devices.
Aims. The aim in terms of ethnomusicology is to explore the expressive cues in a context of a particular musical style and an instrument. The aims in terms of cognitive musicology is to analyze improvised folk music, and to design a novel analysis method for this purpose.
Main contribution. Studies of virtuoso performance of real, partially improvised music have so far been scarce. This study applied a novel method of analysis to explore expressive cues in improvisatory and non-notated musical materials. The consisted of phrase-level segmentation, tempo deviation and frequency-content analysis, and a summary of acoustic and musical features. This analysis discovered the effects that the player employed to convey expression on this particular instrument in this style. These features, roughness of sound, high dynamic variation and twangy sound quality may be regarded as idiomatic to gypsy swing.
Implications. An important portion of all music (e.g., popular music, folk music) is performed without any available notation. Existing analysis methods do not cover such renditions particularly well. In order to gain understanding on expression in musical performance, an anthropologist needs to acquaint herself with the pertinent methods of cognitive musicology.

Keywords: Expressive performance, gypsy swing guitar style, music analysis, acoustic features, improvisation

•*Correspondence:* Marko Aho, Nuupalankatu 15, FI-38210 Sastamala;
tel: +358405677424, e-mail: marko.aho@uta.fi
• *Received*: 30 August 2011; Revised: 16 January 2013; Accepted: 20 January 2013
• *Available online*: 01 February 2013
• doi: 10.4407/jims.2012.11.001

Introduction

Documented instances of Roma individuals playing lutes or plucked instruments at the royal courts of Europe go back over 500 years (Kertesz-Wilkinson, 2001, p. 614). Genres which embody gypsy music in the mind of today's audience – Spanish flamenco, Hungarian primás-tradition, Romanian taraf-orchestras as well as the genre of Russian gypsy romance – were developed in urban surroundings during the 18th through early 20th centuries. It was in association with these and other coexisting indigenous styles the term 'gypsy music' acquired its' broader meaning, as the gypsy image brought exotic flavour to the arts of the romantic era. The gypsy entertainers occupied a special place in the imagination of the European public. Franz Liszt, a keen fan of gypsy music and musicians, stated that when a gypsy fiddler improvised on a tune, he was able to produce the kind of music that synthesized the rational and the irrational, *technical competence* and *depth of feeling* (Malvinni, 2004, p. x).

The virtues of gypsy music find their magnificent exponent in the gypsy jazz style developed in France in the 1930s. The *gypsy jazz* or *gypsy swing* idiom (sometimes also called *manouche jazz* due to its French origins) was created largely, although not single-handedly, by guitarist Jean "Django" Reinhardt. Despite the fact that swing as a form of popular music and jazz fell into obscurity already at mid-century, Django's legacy still lives on, and Django Reinhardt's approach continues to form the basis for the surviving gypsy jazz guitar style in the present day. The local styles of the different gypsy tribes are in a global context being exchanged for a more generic gypsy swing guitar style, which however, retains the basic principles of performance on a particular design of steel string guitar conceived by the Italian luthier Mario Maccaferri in the 1930s. Indeed, the Maccaferri guitar type, originally produced by the French company Selmer, with its' unique features, is a central factor in the stylistic-tactile execution of the style.

In has long been noted by some ethnomusicologists (and comparative musicologists before that) that there exists a direct link between the structure of music and the musical instruments involved in the production of music. Ethnomusicologist John Blacking, in his study on the Nsenga kalimba and the Butembo flute (1955), suggested that the shape of the music was influenced by the spatial properties of the instrument. Curt Sachs (1962, p. 110) wrote: "The instrumental impulse is not a melody in a 'melodious' sense, but an agile movement of the hands." John Baily (1985), in turn, showed how the music of the dutar, a stringed instrument from Afghanistan, was adopted to the performance of music associated originally with another regional plucked lute, and changed the music according to Dutar's physical characteristics in the process. Given that, at least in the cases of acoustic and electro-acoustic musical instruments, musical performance involves a direct relationship of the performing body and the musical instrument's inherent sonic, dimensional, or in short tactile features, it may be argued that also the *expressive conventions* of a musical style are rooted to the possibilities allowed by the musical instruments idiomatic to these styles.

In music psychology and cognitive musicology, the expressive characteristics of music have received considerable attention since the days of the pioneering work of Carl E. Seashore (1938), and led to the establishment of a systematic research tradition in the performance studies and expression during the 1990s. An abundance of research has focused on expressive parameters in *piano performances* due to conveniently measurable properties of the instrument (e.g., Disclaviers producing accurate measurements in MIDI) and easy availability of the recordings of the same works (e.g., Repp, 1990). Whilst most of these studies focus on psychological or technical aspects of the expressive performance cues, musicological studies of performance using contemporary methods have received more attention during the last decade (e.g., Rink, 1995, 2002; Clarke, 2004; Cook, 2004). Another trend in the performance studies has been the tendency to collect the data in real performance situations (Clarke, 2005).

Recent computational models of expressive performance have extended the underlying representation into audio (Widmer & Goebl, 2004; De Poli, 2004) and have extended the choice of instruments beyond the piano (Leman, Desmet, Styns, van Noorden & Moelants, 2009). Interestingly, although guitar was in fact one of the instruments that received emphasis on early studies of emotional performance (e.g., Juslin, 1997; 2000; Erkut, Välimäki, Karjalainen & Laurson, 2000), it has received little attention since these explorations. Yet, guitar in its many variations provides a rich source of cues for expressive performance studies. Also, the majority of the research on musical expression has dealt with performance of Western tonal art music and improvised folk music has received little, if any, attention. This gap in the research area is undoubtedly related to the lack of notation of the improvised folk music. For this reason, novel methods of analysis need to be formulated.

The gypsy swing guitar style seems like a fruitful context for testing such a method. The stylistic and performance elements of the gypsy swing, virtuosity, speed, dramatic gestures such as rapid diminished arpeggios, together with the astonishing dynamic variation allowed by the Selmer-Maccaferri guitar construction, appear as a perfect recipe for a depth of expressive delivery. The internationally famous performers of the style, such as Bireli Lagrene, Stochelo Rosenberg or Joscho Stephan, are indeed all known for a breath-taking display of virtuosity and fierce emotion. For this study, one such internationally prominent artist was willing to demonstrate the devices used for the effect of expressive playing in gypsy guitar style. Here we provide a sample of musical expression on guitar, which combines style known for its expressive nature and an instrument which allows for a flexible tactile manipulation of sound, and a world-class performer.

The Selmer-Maccaferri guitar and the execution of gypsy swing style of playing

An essential component of the highly expressive delivery in gypsy swing is the type of guitar used in this style of music. The construction was designed by the Italian artist and luthier Mario Maccaferri during the swing era of pre II WW. The design was meant as an alternative to the American archtop-construction with the objective of achieving a guitar sound with exceptional “cutting power” to manage the context of jazz bands of the time. In the pre-amplifier era the use of guitar as a solo instrument on a jazz band context was very limited.

Figure 1. A copy of a Selmer Modelé Jazz guitar.

The guitar construction that Maccaferri designed for Selmer possesses a unique bright and cutting, non-linear tone. The construction of a high quality Maccaferri type guitar is light with laminated back and sides and a minimum of inner bracing allowed by the slightly domed shape of the soundboard (see Figure 1). The strings are of low gauge, and they are set on a distinctly high action compared to other types of flat top steel string guitars. For full effect, the inherent features of the guitar must be combined with a special picking technique; the pick itself must be very thick and hard (Django himself used e.g., large buttons), and most of the picking is downstrokes aiming at a strong string vibration adversely to the plane of the soundboard, as this produces the soundboard to vibrate strongly and thus a high volume. Lee, Chaigne, Smith & Arkas (2007) made a study on the characteristic acoustic functioning of the Selmer-Maccaferri type guitar. The distinctive heavy downstroke picking is the cause of a pitch shift heard in the sound. Initially the pick displaces and releases the string quickly on the horizontal plane. Lee et al. were able to observe that the fundamental at the onset is approximately 5Hz higher than the steady-state frequency. Furthermore, the horizontal plane is vibrating at a slightly higher frequency than that of the vertical plane even after the initial stretching of the string from a particularly pronounced attack of a pick. According to Lee et al., the bright “twangy” sound of the guitar results for the most part from the top-plate which resonates high frequencies greater than those of the more rigid top plate of a classical guitars, the plot of the driving-point impedance of the former taking place between 400 to 5000Hz.

A strong and fast vibrato, which also enhances the sustain of the note, is a final trademark of gypsy swing artists. The low gauge strings together with the thin and thus responsive top plate allow for a wide variation in the manipulation possibilities of the string acoustics in playing: the guitar allows very wide dynamics; and, the overall non-linear quality of the acoustic sound of the guitar gives a lively and interesting character to the sound by default. These qualities make the gypsy swing guitar a powerful tool of expressive playing in capable hands.

Expressive performance

An expressive performance is influenced by both the structure of a music composition (or some other discrete set of pitches) and by the ways these structures are brought out by the performer, both contributing in their own particular way to the emotional meaning of music (Thompson 2009, p. 144). The study of performance has been, by and large, study of expression, and this expression has been predominantly associated with emotional expression (Gabrielsson 2003, p. 231). The studies have mainly focused on two aspects of the communicative process: the accuracy of the communication and the codes used by performers and listeners. The studies have indeed confirmed that professional music performers are more or less able to communicate emotions to listeners (for review, see Juslin & Timmers 2010, p. 453-492). The research has shown that a variety of codes are in play here. Repp (1999) reported extensive measurements of timing and dynamics in piano performance. It turned out that timing and dynamics, which were thought to be the most expressive performance variables, accounted for only a small part (10-18%) of the 'overall aesthetic quality' ratings of expert judges.

Indeed, already Seashore (1938, p. 24) was convinced that *timbre* plays an important role in the expression. In a study of clarinet performance expressiveness, Barthet, Depalle, Kronland-Martinet, and Ystad (2010) were able to demonstrate that timbre, as well as timing and dynamics mediate expressiveness. Holmes (2011), in an interview study of an elite guitarist, emphasised the role of timbre at the heart of expressive performance. It seems that the difficulty of measuring timbre (Scherer & Zentner, 2001), and the technical means by which timbre is created (Holmes, 2011, p. 2), have kept this element of music a late arrival in the study of expression in music.

Based on the accumulated research made in the 1990s, Juslin (2003) made an attempt to create a comprehensive model of expression in music, namely the GERM-model. This model takes into account a) musical structure; b) emotional expression; c) human random variations; and d) biological movement principles. Given that all these factors take place simultaneously in a musical performance, there are substantial challenges to the study of musical expression.

Aims of the study

Previous studies on expressive performances have predominantly been based either on a real performances with no neutral point of comparison (i.e., most studies on expressivity in piano performance, see e.g., Widmer & Goebl, 2004; Goebl & Palmer, 2009) or comparison of various emotional expressions to neutral (e.g., expressing well-known tunes such as Frere Jacques or When the Saints) using different emotional expressions (e.g., Juslin, 1997, 2000), or focusing on ornamentation of keyboard performances (Timmers & Ashley, 2007). However, what has *not* been studied, is virtuoso performance of real, partially improvised music.

Therefore the first aim is to develop a methodology to study expression in improvisations in order to explore the devices that a competent gypsy swing guitarist uses to achieve an emotional expression. This study presents a novel challenge for the expressive performance analysis, since there is no standard point of comparison, such as the notated score. The actual musical contents may also vary across performances. For this reason, we outline a novel method of analysis suitable for improvisatory and non-notated musical materials. Our second aim is to discover the acoustic and musical features of expressive performance produced by a master gypsy swing guitarist in a real, albeit purposefully arranged playing situation.

The focus here is on the expressive performance cues within a particular musical style, since the idioms of the style in question are very much a result of the tactile and sonic potential of a particular musical instrument. Finally, this study deviates from the past research on musical expression which has dealt with performance of Western tonal art music. Folk music and improvised folk music have not been explored within the expressive performance tradition research.

Expression analysis scheme for improvised music

Materials

A competent gypsy swing guitarist was asked to participate in a study involving expressive performance. This musician is an internationally touring recording artist, who has featured at the Samois-Sur-Seine festival, the major event of the international gypsy swing scene. The guitar he used in the performing the improvisations was a high-end Selmer-Maccaferri type guitar of the brand AJL Guitars. He was asked to perform several pairs of impromptu choruses of common jazz tunes. In both versions, he was advised to play according to his professional level. He was asked to perform one version in a purposefully non-expressive manner and second version with as much expression as he could.

The musician chose to perform first *Nuages*, a composition by Django Reinhardt. We limit our analysis on the Nuages because it is an original composition of Django Reinhardt, the iconic gypsy swing master, and was performed in a rubato solo manner which can be heard on many recordings by gypsy swing artists such as Bireli Lagrene or Django Reinhardt himself, or in the intros of countless slow or medium-tempo gypsy swing performances. The recording was done in a relaxed setting. The player was sitting on a couch, the recording equipment was on the background and the recording proceeded according to his own pace. The recording was captured to a portable computer using a Samson CO1U microphone to 44,1 kHz 16-bit PCM wave files. Sound examples of the performances are available at http://www.xxx.fi

The outline of the analysis

Since we were analysing an improvisation, there was no score or actual target notes that could be used as a point of comparison. For this reason, we needed to establish novel ways of mapping the two improvisations differing in their expressive quality into each other. We think this is an important direction to be made in general in the performance analysis, since an important portion of all music (e.g., popular music, folk music) is performed without any strict notation at hand. We outlined three steps that are necessary for the comparison of such performances. These are (1) phrase-level segmentation verified by tempo deviation analysis and frequency-content analysis, (2) extraction of features related to expression, and (3) the comparison of the acoustical and musical features from each of the identified expressive and non-expressive phrases.

Phrase-level segmentation

First, a manual segmentation of the performances into phrases was carried out. The purpose of this was to ensure that the comparison is carried out within meaningful units (phrases are assigned alphabets A-I in Figure 2). Similarly marked sections in the upper and lower panels of Figure 2 readily suggest how different the tempi and dynamics are involved in the expressive and non-expressive renditions of the same improvisation. In some cases (notable in the phrase D, G, and I), the performer adds variations and extra notes, arpeggios or chords to the phrase, so they are by no means identical in the two versions.

Since the two performances contain different temporal deviations, with the assumption that the expressive performance would have more pronounced rubato characteristics, we calculated the optimal temporal alignment of each phrase by means of dynamic mapping provided by *Match Performance Aligner* (Dixon, 2005). The result yields a ratio of tempi between expressive and non-expressive performance for each phrase. These are shown in the lowest panel of Figure 2. If the tempi of the two performances are identical, the ratio is exactly 1 throughout the phrase. If the values are above 1, expressive performance is slower than the non-expressive, and when the values are lower, the expressive performance is faster in comparison with non-expressive performance. Most of the deviations result from rubato differences but certain large peaks in the tempo ratios result from the two phrases having different musical material in them. For instance, the end of phrase F has a substantial indication of expressive performance slowing down, which does not actually happen, but the performances have different passages at the end of the phrase. An example of a very similar timing across performances is the phrase E, in which the tempo ratio is constantly close to 1.

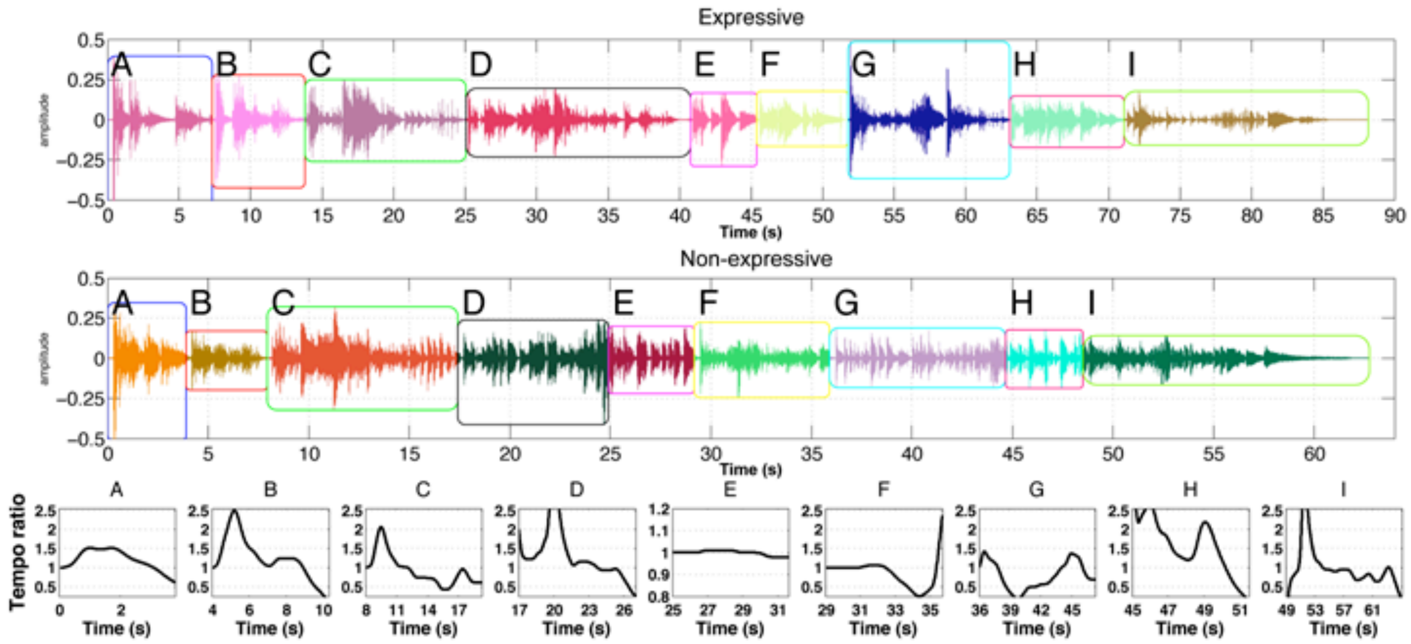

Figure 2. Segmentation of the performance into phrases. The two upper panels show the corresponding phrases between the two versions (expressive and non-expressive) and the lowest panel displays tempo deviation between the two.

Since the two performances do occasionally have different musical contents, not only expressive dynamics and temporal variations, it might be prudent to visualize the actual frequency contents of the phrases. Again, no official score notation is available and resorting to one would be against our arguments that expressive and improvisational qualities of music can be investigated without resorting to a strict deadpan template.

The fundamental frequencies of detected onsets in all phrases were estimated using autocorrelation-based pitch estimation, based on Tolonen and Karjalainen (2000), with multi-pitch capacity, implemented to MIR toolbox (Lartillot, Toiviainen, & Eerola, 2008). The estimated pitch range was constrained to 75 to 1200 Hz in order to improve legibility of the ensuing figures, although this leaves out certain high frequency components. Figure 3 shows the patterns delineated by each phrase for each performance (note that the Y-axis is logarithmic to convey more score-like visual representation for the reader).

Some of the softer onsets (e.g., the onsets between seconds 2.5 and 4 in phrase A) are not registered in the pitch estimation due to the method involved, which also helps to increase clarity and allow the main patterns of each phrase to be discernible. Also, the most performed harmonics, fretboard sounds, vibrato and other expressive devices cannot be seen in the pitch estimations either due to threshold related to onset detection or pitch estimation. Displaying the full details with carefully chosen spectrograms would be another option but in our opinion the pitch tracks of the guitar performances are in this case more illustrative.

Despite these representational limitations, the overall pitch patterns of the phrases are evident in Figure 3 and exemplify the wide range of performance options available even when performing basically the same tune in repeated fashion. In most phrases (e.g., A, B, C, G, H), there is a high resemblance between the two performances in the pitch tracks although, in some cases, the frequency contents of the phrases appear to be discrepant (D, E, F, and I). A closer look and aural investigation of these phrases reveal that whilst the initial melody line is mostly identical, heavy variation in terms of virtuoso scalar and arpeggio patterns create a large number of permutations and pitch candidates, which are challenging for individual pitch estimation. For instance, phrase E is a typical example in which the non-expressive version has actually more elaborate ornamentation around the basic chords. Also, the chord arpeggios passages (both up and down) are easily visible in pitch estimations as ascending and descending patterns (particularly phrases F, I and G).

At this stage, we have verified that the analysis of expressive features is based on comparable units (phrases) despite the heavy improvisatory ornamentation and variation present in music. In the next section, we will look at detail what kind of acoustic and musical features set the two performances apart.

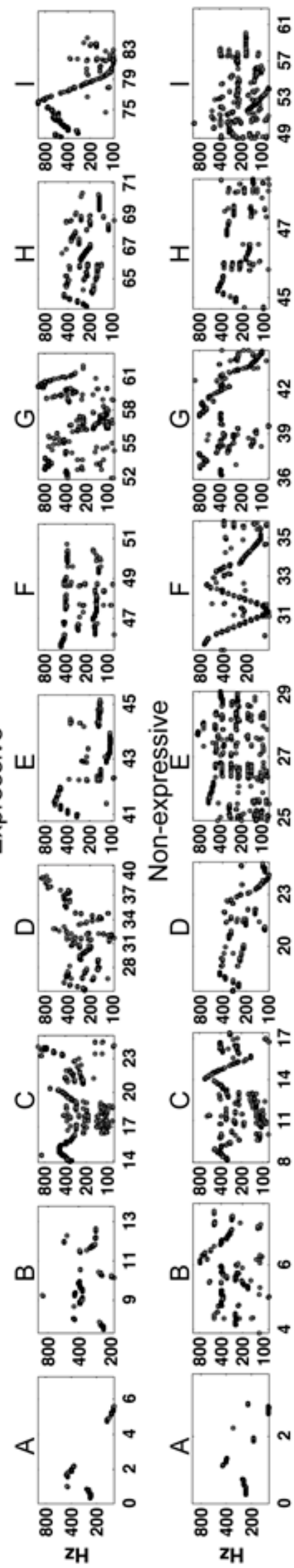

Figure 3. The pitch patterns delineated by each phrase for each performance showing the overall similarities of the identified phrases.

Extraction of features related to expression

A more detailed analysis was carried by extracting a set of features from each phrase (A-I). We chose a handful of features known to be important expressive characteristics in piano and violin performances, such as *dynamics* (amplitude and amplitude variation, see e.g., refs here), *articulation* (attack time and fluctuation, see refs), *tempo*, and timbre (spectral centroid, see for example ref). We were also curious about some alternative formulations of these features?, such as roughness and spectral novelty as a measure of timbre and pulse strength as a measure of expression, since these are known to be important in conveying emotions in music (see Eerola, Lartillot & Toiviainen, 2009; Eerola, 2011). However, we did not want to have an excessive long list of features since that would provide both practical and statistical problems for the analysis. To safeguard against too many similar (intercorrelating) features, we calculated the absolute mean correlations between the nine features on this dataset, and the differences were not significant (r=.20, p=ns). In sum, the chosen features are predominantly representing individual musical factors. Table 1 offers a summary of all computationally extracted features.

Table 1. Summary of computationally extracted features.

Feature	Explanation
RMS (M & SD)	Root Mean Square energy of the signal, which indicates the amplitude of the sound volume (M=mean, SD=standard deviation, which is used to describe the dynamic change of the energy).
Attack time	What is the time duration between the initial note onset time and its peak time, i.e., how percussive or soft is the sound attack. The signal envelope is used to estimate the peaks and valleys needed.
Fluctuation	The prevalent rhythmic periodicity across a wide range of periods (Pampalk et al., 2002).
Tempo	Estimation of tempo (in BPM) by looking at transformation of the amplitude envelope using autocorrelation-based technique to estimate the prevalent periods. An emphasis on perceptually salient periods is used (Toiviainen & Snyder, 2002).
Pulse clarity	How clear and stable is the pulse in music, also called beat strength. The model, proposed by Lartillot and his colleagues (2008), uses onset characteristics and the periodic nature of onset trains to estimate the clarity of pulse.
Roughness	How high is the sensory dissonance of the sound. In this measure, the peaks in the spectrogram are utilized to estimate the roughness, which is based on a psychoacoustic model by Sethares (1998).
Spectral centroid	What is the geometric centre of the spectral frequency. A high value indicates prevalence of high-frequencies, which makes the sound more sharp and less soft (Juslin & Laukka, 2003).
Spectral novelty	Measures how dissimilar the spectral content is over time. The lower the self-similarity of the spectrum, the more novelty it can be said to contain. The measure is based on detection of edges within the diagonal of the self-similarity matrix (Foote & Cooper, 2003).

All features were analysed from each phrase with an equal number of dynamic analysis windows. This was done in order to get estimates of the variance of the features, since using single windows equal to the length of the phrases would mostly have yielded a singular estimates and using a constant window length for this would have provided a considerable variance – and hence reliability – in terms of the number of observations per phrase since they differ in length. All features were extracted with the *MIR toolbox* (Lartillot, Toiviainen, & Eerola, 2008, version 1.3.3) using a frame-based approach. For roughness, which needs to be analysed using short-term windows, analysis frame of 120 ms was used with a 50% overlap between frames. For spectral novelty, the frame length was 100 ms with 50% overlap. The results from the frames were then summarized by the mean and the standard deviation. As all the features are available and documented in MIR toolbox[1], we will not explain all their technical details in full here.

The comparison of the acoustical and musical features in expressive and non-expressive performances

Figure 4 shows a visualization of the nine chosen features means and variances across the nine phrases. Each point in time represent one phrase (labelled with alphabets) and each panel contains results from one feature for both expressive (white) and non-expressive (black) performances. In many cases the performances do not differ greatly in terms of the chosen features, but certain ones such as RMS, tempo, rhythmic fluctuation, pulse clarity, roughness and particularly spectral novelty seem to be picking up the expressive variations. In certain cases, spectral centroid for instance, large differences are evident but only in few phrases (in this, in the phrases F and I). There are also interdependencies between the features, which are interesting both from musical and acoustic point of views. For instance, the expressive performance has lower dynamics, that is, it is often more softly played which seem to reflect over to higher spectral centroid (brighter sound), and lower roughness in general. Expressively performed phrases are also slow but this leads to a communication of a more clear pulse (higher pulse clarity on average). The defining characteristic is also richer sound variation, as indicated by consistently higher spectral novelty of the expressive performance in comparison with non-expressive ones. To check whether this was due to performance 'artefacts' (i.e., fretboard noise), the phrases were checked separately but it seems that the expressive performance is simply more diverse in its sound quality. Finally, we need to point out can that the expressive characteristics are not similar across all phrases, an issue which will be examined in the next section.

[1]See https://www.jyu.fi/hum/laitokset/musiikki/en/research/coe/materials/mirtoolbox

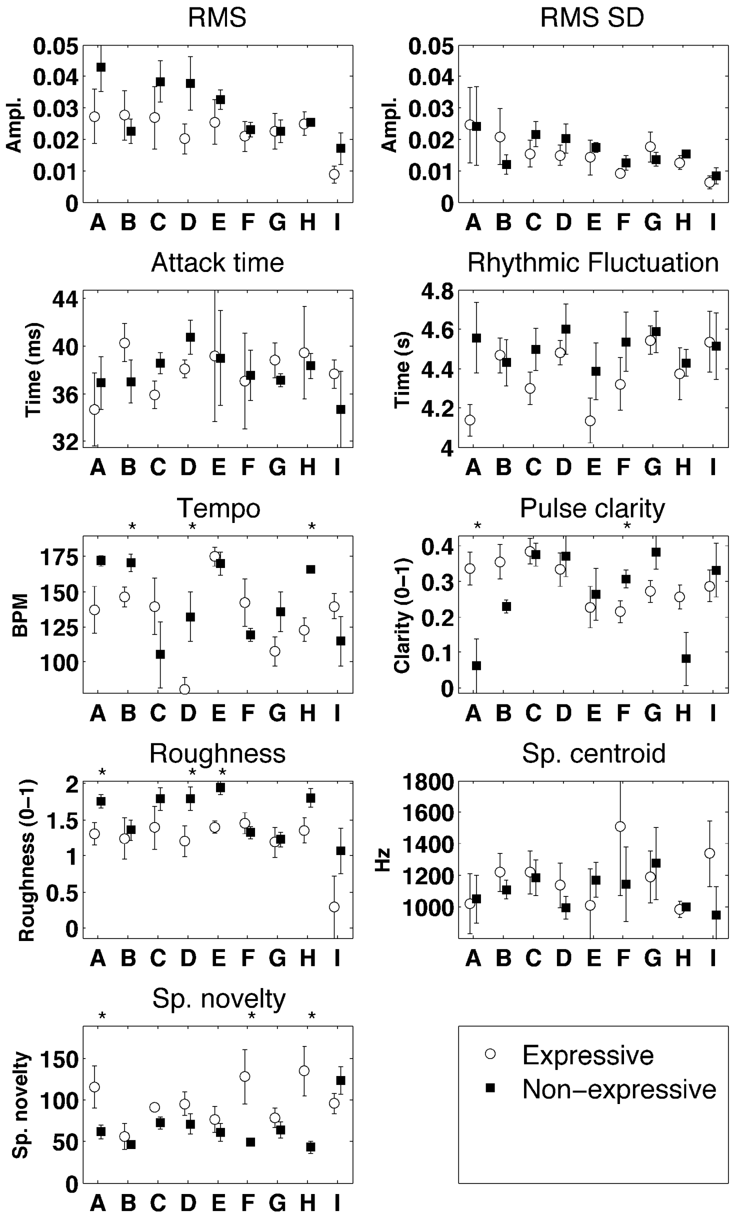

Figure 4. Visualization of the nine chosen features means and variances across the nine phrases.

To examine the above-mentioned differences more rigorously, we first ran a Kruskal-Wallis analysis across observations obtained from each of the nine phrases, and this analysis was carried out for each feature separately. This particular choice of analysis of variance was motivated by the small amount of observations (5 samples from each phrase, yielding 45 observations in total for both performances) producing a rather small sample and a non-normal distribution for a number of variables (checked with Lilliefors test for variable normality at the p<.05 level). The results of the analysis are displayed in Table 2 (column Expression). From this we learn that the dynamics (RMS, X^2=4.76, p<.01) are significantly higher in the non-expressive performances (M=0.029) than in expressive performances (M=0.023). Also the rhythmic fluctuations are reliably different across the two performances (X^2=5.78, p<.05), in which the non-expressive typically has higher fluctuations than the expressive performance (means of 4.3 and 4.5). Since the rhythmic fluctuation is an index of the prevalent periodic patterns, we interpret this to reflect often the higher repetition rate of the melodic patterns, also reflected in the tempi. Roughness was higher in the non-expressively performed phrases (X^2=10.58, p<.01), which might reflect the fact that they were also typically performed somewhat louder (higher RMS) and the increase in dynamics for this instrument increases the harmonics and the twangy noise components at the note onsets, which both give rise to higher sensory dissonance. Finally, the spectral novelty was markedly higher (X^2=13.90, p<.001) in the expressive performances, suggesting clearly that the sound quality is produced in significantly richer fashion in this performance than in the non-expressive one.

Results of the analysis

It could be characterized from this analysis that the expressive performance is softer, brighter, less dissonant, yet richer in its temporal and sound structure than the non-expressive performance. Indeed such a description fits with the aural image one gets from hearing the two performances. We have to keep in mind that the non-expressive performance is played on the same instrument by the same performer, and known to contain traces of the expressive performance devices (e.g., Woody, 2003).

Table 2. Summary of differences in acoustic and musical features across all phrases.

Feature	***Expression***	**Phrase**	**Interaction**
RMS	p<.05	p<.05	-
RMS SD	ns	Ns	-
Attack time	ns	Ns	-
Rhythmic fluctuation	p<.05	Ns	-
Tempo	ns	p<.001	p<.01
Pulse clarity	ns	p<.001	p<.01
Roughness	p<.01	p<.001	-
Spectral centroid	ns	Ns	-
Spectral novelty	p<.001	p<.05	p<.01

Note: df = 1,88 for the expression and df = 1,78 for the phrases.

Aside from the global inspection of the phrases, certain phrases do deviate from the overall pattern of the results in Figure 4 (e.g., phrase B in RMS, G in pulse clarity and so on). This suggests that there could be different expressive cues for different phrases. This idea was formally tested by a two-way ANOVA with Expression and Phrase as the main factors to explore whether there are differences across the phrases and if these differences are mutually interacting. Table 2 shows the results for the phrases and also the two interactions between the Expressions and Phrases. RMS values do vary significantly across the phrases, particularly between the first and the last phrase in both performances. Tempo, pulse clarity, roughness and spectral novelty also differ between the phrases, which makes sense, since the phrases have entirely different musical content in them. Those phrases which significantly differ from each other have been identified for each feature in Figure 4 with small asterisks. Note that these results are not dependent on the overall length of the phrases, since both performances are analysed within similar sub-divisions of the phrases. The only interactions between the expressive performances and phrases occur with tempo and pulse clarity. This means that these features do exhibit different outcomes depending on both the performances and the phases. For instance, in the first two phrases, the tempo is higher for non-expressive performance whereas for the C and F I phrases, the expressive performances are faster. In the remainder of the features, no such interactions exist.

Discussion

It is often argued that a successful musical performance in general must contain both technical and expressive aspects (e.g. Gabrielsson 1999; Sloboda 1996). Still, some forms of music (e.g. electronic music) no doubt involve emotional expression of the performer in lesser extent than others (e.g. many forms of Afro-American popular music). Some musicians are better than others in conveying emotions. Finally, different instruments allow for different expressive cues, and some musical instruments seem to allow more expression than others (Gabrielsson & Juslin 1996, p. 87).

The expressive performance differed from less-expressive performance by number of subtle musical cues, namely by timbre (roughness and spectral novelty), dynamics, and by overall pace of the rhythmic fluctuations. The expressive cues did not, however, remain static during the whole performance, but varied dynamically across the phrases. This was clearly evident in the analysis, although at times it is difficult to disentangle the differences related to separate expressive characteristics and musical contents since the two are always interrelated in real improvisations. It must be kept in mind that the comparison in this case was done without a strict manual annotation that allows us to focus reliably on attack parts and steady state parts of the sounds. This might have been technically possible and feasible for the analysis, and it seems a task worth undertaking to develop next stage in the analysis of improvised music. Such a method must be able to handle large amount of audio material in order to be lucrative for more comprehensive research projects so we chose not to follow the road to manual annotation. It must also be noted that the array of musical and acoustic was by no means complete and the features used in the present study can be implemented in different ways. We could also have included various timbral descriptors such as jitter, vibrato, shimmer and inharmonicity. However, the number of features was considerably large for an explorative analysis and an increase in the number of features would have risked an increase in false positives in this study design. Finally, this novel analysis method of expression in improvisation proved to be successful at least in the technical sense and provided interpretable insights into the expressive devices utilized in this case. Whether such preliminary results can be extended onto a larger sample of improvised music, remains to be seen.

The object of study, musical expression, was expression of an individual, but an individual expressing on a special type of instrument, an instrument which is an underlying cornerstone of a particular musical style. Gypsy swing is primarily played in an ensemble with a steady beat, but rubato solo performances are not rare either. Instrument ergonomics influence the structure of music, but the tactile features of the instrument also influence the expressive conventions of a musical style. In view of the formal analysis made we could detect, that the player indeed was able to utilize the particularities of the Selmer-Maccaferri guitar type to great effect. These features, roughness of sound, high dynamic variation and general twangy sound quality may be regarded as conventional or idiomatic to gypsy swing. As to which features would result in enhanced expression, the analysis showed maybe slightly surprisingly, that expressive playing had lower dynamics, is softer and comes with lower roughness and dissonance than non-expressive playing.

If one tries to avoid disagreeing with C.E. Seashore in that musical art generally relies on "artistic deviation from the fixed and regular: from rigid pitch, uniform intensity, fixed rhythm, pure tone and perfect harmony" (Seashore 1938, 29), one must conclude that a prominent expressive trademark of gypsy swing is the *highly rich variation of sound quality*. Alongside with variation in the temporal structure, which is firstly more or less applicable with musical instruments in general and secondly not really possible in the typical gypsy swing ensemble playing, spectral novelty was the only measured feature which would result in more expressive intent with increase. Then again, both early and recent studies have emphasized the importance of timbre to expression (Seashore, 1938, p. 24; Repp, 1999; Barthet, Depalle, Kronland-Martinet, & Ystad, 2010; Holmes, 2011). Indeed, the Selmer-Maccaferri guitar really shines on how sensitive it is to slight variation of the plectrum touch. In this sense, an archtop jazz guitar, which produces a much more uniform tone throughout its dynamic range is no match. On a Selmer-Maccaferri type guitar, a forceful down-stroke on a "sweet spot" of the string resulting in the initial pitch-shift and highly un-linear sound is just one end of a continuum – the player has a whole range of sounds available from soft and full to hard and twangy.

The analysis of improvised, non-notated expressive performance provides a real challenge for formal analysis, since the existing methods do not cover such renditions particularly well. Here the problem was approached by outlining a simple analysis strategy with a definition of units of analysis (phrases), corroborating that the temporal and pitch contents roughly match, and then a number of computationally extracted features were obtained for each phrase. These features provided multiple ways of discriminating the expressive characteristics from the non-expressive ones, despite the lack of real deadpan performance or other commonly used strategies for such estimation. Although the sound analysis made in this study was revealing, it may be worth pondering upon whether this kind of study of improvised, masterful performance could make use of some kind of phenomenological methods as proposed by Gabrielsson (1995) in order to have valid interpretations of the features which lend themselves to measurement. More and more distinctive means of measurement audio data have been developed since mid 1990's, but already then Gabrielsson and Juslin (1996) noted that "it has become increasingly clear that the main problem is rather to find ways of interpreting the wealth of performance data in a meaningful way" rather than acquire more data. The excerpts used in this study could possibly be listened to analytically by a number of people in order to obtain verbal descriptions of what is happening moment by moment in terms of expression. In this way, some interesting knowledge might be obtained from, for example, vibrato, which was now excluded from analysis this time in effort to keep the parameters to be analyzed at a manageable number. Different kinds of vibratos seem to have differing expressive identities (see Gabrielsson & Juslin, 1996).

There is no doubt that expression is likely to be affected by the presence or absence of audience, and also by the performer's momentary disposition and mood. The air of the initial recording situation was that of professional nonchalance, almost indifference, but in retrospect it seems obvious that the presence of one of the authors and the recording context must have provided at least some stimulation for the artist. It seemed he did not reflect much upon his task of playing emotionally empty and charged pieces of improvisation, but proceeded on gut feeling and relied on his intuition. He did not seem to expect any feedback on his performance between takes. The more refined gestures and expressive achievements tell of an exceptionally able player, who can probably use contrasting and even conflicting expressive devices to his artistic means (e.g. to play an otherwise happy tone in a sad expression may transcendent the expression to a totally new level). The expression is affected by the instrument and the style, which in this case are exceptionally interdependent, but ultimately we are dealing with the emotionally expressive fingerprint of the player playing that instrument in a particular style.

To what extent do the expressive conventions of a musical style reflect the general mentality of that style? Earlier on we cited some historical descriptions of gypsy music: in general it has become associated with highly expressive and virtuosic delivery. This study hopefully managed to shed some light on the other one of these components, expression, which is certainly involved in the recent global success of the gypsy swing guitar. The Selmer-Maccaferri guitar construction, and the forceful way it is supposed to be played to get the best out of it, almost automatically results in a more or less noisy attack and rough timbre, which are apparent in both expressive and non-expressive performances. A pioneering study made by Gabrielsson and Juslin (1996) employed six guitarists playing electric guitars. Gabrielsson and Juslin had the players perform several discrete emotions, and then had these performances rated and analysed them with regard to their physical characteristics. With regards to guitar, Gabrielsson and Juslin were able to conclude e.g. that "angry" tone was produced with a combination of a hard, "noisy" plectrum attack and strong upper partials, or "rough" timbre; "sad" versions were played with a deep and slow vibrato, whereas the "fearful" versions were played with a very fast but more shallow vibrato with some irregularity; "happy" versions used a fast and very light vibrato. All in all gypsy swing appears as both angry and happy. In ensemble settings, the tempi on which gypsy swing performance are delivered are often particularly fast, as in both angry and happy music the tempi should be. The favoured jazz standards (old Tin-Pan-Alley tunes) are, generally speaking, happy tones in major modality. The fast and light vibrato indicates happiness. Is the mentality of gypsy swing, then, happy, with a tint of aggression? Let us add that the other central feature of gypsy swing, which was not discussed here but which is well worth further study, is the collection of lightning-fast virtuoso gestures, which come in bursts, hits and sudden accelerations, to lend a few terms conforming to Daniel Stern's (Stern 1986) theory of vitality gestures. Is gypsy swing and are its' proponents essentially happy but potentially dangerous? This tentative question can of course ultimately be answered only by a study of the gypsy swing as a cultural idiom.

The two other popular melody instruments of Middle-European gypsy music, the fiddle and the clarinet, offer similar (but also different) ways of expression compared to guitar. An interesting detail is that Django Reinhardt's two most important band-mates and co-soloists of his heydays, Stéphane Grappelli and Hubert Rostaing, played the violin and clarinet respectfully. Although the present formal analysis was purely exploratory, the analysis scheme is able to facilitate an analysis of a significantly larger number of music excerpts – possibly from different performers and instruments of a common style.

References

Baily, J. (1985). Musical Structure and Human Movement. In P. Howell & al. (Eds.), *Music Structure and Cognition* (pp. 237-258). London: Academic Press.

Baily, J. (1992): Music Performance, Motor Structure, and Cognitive Models". In M. P. Baumann, A. Simon and U. Wegner (Eds.), *European Studies in Ethnomusicology: Historical Developments and Recent Trends* (pp. 142-158). Berlin: International Institute for Comparative Music Studies and Documentation.

Barthet, M., Depalle, P., Kronland-Martinet, R., & Ystad, S. (2010). Acoustical Correlates of Timbre and Expressiveness in Clarinet Performance. *Music Perception, 28* (2), 135-154.

Blacking, J. (1955). Eight Flute Tunes from Butembo. *African Music, 1(2),* 24-51.

Charle, F. (1999 [2008]). The Story of Selmer Maccaferri Guitars. Paris: Francois Charle.

Clarke, E. F. (2004). Empirical methods in the study of performance. In E. F. Clarke and N. Cook (Eds.), *Empirical Musicology. Aims, Methods, and Prospects* (pp 77–102). University Press, Oxford.

Clarke, E. F. (2005). Creativity in performance. *Musicae Scientiae, 9*, 157-182.

Clynes, M. (1977). Sentics. The Touch of Emotions. Garden City, New York: Anchor Press/Doubleday.

Cook. N. (2004). Computational and comparative Musicology. In E. F. Clarke and N. Cook (Eds.), *Empirical Musicology. Aims, Methods, and Prospects* (pp. 103-126). University Press: Oxford.

Cruickshank, I. (1985). *The Guitar Style of Django Reinhardt & the Gypsies*. London: Wise Publications.

De Poli, G. (2004). Methodologies for expressiveness modelling of and for music performance. *Journal of New Music Research, 33*(3), 189-202.

Dixon, S. (2005). Live tracking of musical performances using on-line time warping. *Proceedings of the 8th International Conference on Digital Audio Effects* (pp. 92-97).

Dregni, M., Antonietto, A., & Legrand, A. (2006). Django Reinhardt and the illustrated history of gypsy jazz. Denver: Speck Press.

Erkut, C. Välimäki, V., Karjalainen, M., & Laurson, M. (2000). Extraction of physical and expressive parameters for model-based sound synthesis of the classical guitar. *Proceedings of the 108th AES Convention*.

Foote, J., & Cooper, M. (2003). Media segmentation using self-similarity decomposition. *Proceedings of SPIE storage and retrieval for multimedia databases*, 5021, 167–175.

Gabrielsson, A. (1995). Expressive intention and performance. In: R. Steinberg (Ed.), *Music and the mind machine. Psychophysiology and psychopathology of the sense of music* (pp. 35-47). Heidelberg: Springer Verlag.

Gabrielsson, A. (2003). Music performance research at the millennium. *Psychology of Music 31*, 221-272.

Gabrielsson, A. & Juslin, P. N. (1996). Emotional Expression in Music
Performance: Between the performers expression and Listeners experience. *Psychology of Music, 24*(1), 68-91.

Goebl, W., & Palmer, C. (2009). Synchronization of timing and motion among performing musicians. *Music Perception, 26,* 427-438.

Holmes, P. A. (2011). An exploration on musical communication through expressive use of timbre: the performers perspective. *Psychology of Music 40(3), 301-323*.

Jensen, K. (2007). Multiple scale music segmentation using rhythm, timbre, and harmony. *EURASIP Journal on Applied Signal Processing, 2007(1),* 159-159.

Juslin, P. N. (1997). Emotional communication in music performance: A functionalist perspective and some data. *Music Perception, 14,* 383-418.

Juslin, P. N. (2000). Cue utilization in communication of emotion in music performance: relating performance to perception. *Journal of Experimental Psychology: Human Perception and Performance, 26*, 1797-1813.

Juslin, P. N. (2003). Five Facets of Musical Expression: A Psychologist's Perspective on Music Performance. *Psychology of Music, 31,* 273-302.

Juslin, P. N., & Laukka, P. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychological Bulletin (129)*, 770-814.

Juslin, P. N. & Timmers, R. (2010). Expression and communication of emotion in music performance. P. N. Juslin & J. Sloboda (Eds.), *Handbook of Music and Emotion: Theory, Research, Applications* (pp. 453-492). Oxford: Oxford University Press.

Kertesz-Wilkinson, I. (2001). Gypsy Music. In S. Sadie (Ed.), *The New Grove Dictionary of Music and Musicians* (pp 613-619). London: McMillan.

Lartillot, O., Toiviainen, P., & Eerola, T. (2008). A Matlab Toolbox for Music Information Retrieval. In Preisach, C. and Burkhardt, H. and Schmidt-Thieme, L. and Decker, R. (Eds.), *Data Analysis, Machine Learning and Applications*, pp. 261-268, Springer, Berlin.

Lee, N., Chaigne, A., Smith, J. O. & Arcas, K (2007). Measuring and understanding the gypsy guitar. *Proceedings of the International Symposium on Musical Acoustics, Barcelona, Spain, September 9-12, 2007* (pp. 1-8).

Leman, M., Desmet, F., Styns, F., van Noorden, L., Moelants, D. (2009). Sharing musical expression through embodied listening: a case study based on Chinese guqin music. *Music Perception, 26*, 263-278.

Malvinni , D. (2004). *The Gypsy Caravan: from Real Roma to Imaginary Gypsies in Western Music and Film*. New York: Routledge.

Mohn, C., Agstaller, H. & Wilker, F-W. (2010). Perception of six basic emotions in music. *Psychology of music, 27*, 1-15.

Pampalk, E., Rauber, A., Merkl, D. (2002). Content-based Organization and Visualization of Music Archives. *ACM Multimedia, Juan-les-Pins, France* (pp. 570-579).

Repp, B. H. (1999a). A microcosm of musical expression: II. Quantitative analysis of pianists' dynamics in the initial measures of Chopin's Etude in E major. *Journal of the Acoustical Society of America, 105*, 1972–1988.

Repp, B. H. (1999b). A microcosm of musical expression: III. Contributions of timing and dynamics to the aesthetic impression of pianists' performances of the initial measures of Chopin's Etude in E major. *Journal of the Acoustical Society of America, 106*, 469–478.

Repp, B.H. (1990). Further perceptual evaluations of pulse microstructure in computer performances of classical piano music. *Music Perception, 8*, 1–33.

Rink, J. (Ed.) (1995). The Practice of Performance: Studies in Musical Interpretation. Cambridge, UK: Cambridge University Press.

Rink, J. (Ed.) (2002). Musical Performance. A Guide to Understanding. Cambridge, UK: Cambridge University Press.

Sachs, C. (1962): The Wellsprings of Music. The Hague: Nijhoff.

Seashore, C. E. (1938). Psychology of Music. New York: McGraw-Hill.

Sethares, W. A. (1998). Tuning, Timbre, Spectrum, Scale. Berlin: Springer-Verlag.

Scherer, K & Zentner, M.R. (2001). Emotional Effects of Music: Production Rules. In P.N. Juslin & J.A. Sloboda (Eds.), *Music and Emotion: Theory and Research* (pp. 361-392). Oxford: Oxford University Press.

Shaffer, L.H. (1981). Performances of Chopin, Bach and Bartòk: Studies in motor programming. *Cognitive Psychology, 13,* 326–376.

Silverman, C. (2000). Rom (Gypsy) Music. In T. Rice, J. Porter and C. Goertzen (Eds.), *The Garland Encyclopedia of World Music* (pp. 270-293), Volume 8 Europe. London: Garland Publishing.

Stern, D. (1986). The interpersonal world of the infant. Basic Books.

Sundberg, J., Friberg, A., & Frydén, L. (1991). Threshold and preference quantities of rules for music performance. *Music Perception, 9,* 71–92.

Thompson, W. F. (2009). *Music, Thought, and Feeling. Understanding the Psychology of Music*. New York: Oxford University Press.

Timmers, R., & Ashley, R. (2007). Emotional ornamentation in performances of a Handel sonata. *Music Perception, 25*(2), 117-134.

Toiviainen, P., & Snyder, J. (2003). Tapping to Bach: Resonance-based modeling of pulse. *Music Perception, 21*(1), 43–80.

Tolonen, T., & Karjalainen, M. (2000). A computationally efficient multipitch analysis model. *IEEE Transactions on Speech and Audio Processing, 8*, 708-716.

Widmer, G., & Goebl, W. (2004). Computational Models of Expressive Music Performance: The State of the Art. *Journal of New Music Research, 33(3)*, 203–216.

Woody, R. H. (2003). Explaining Expressive Performance: Component Cognitive Skills in an Aural Modeling Task. *Journal of Research in Music Education, 51(1)*, 51-63.

Öberg, A. & Horowitz, M (2005). *Gypsy Fire. A Collection of Gypsy Jazz Patterns and Arpeggios for Guitar*. DjangoBooks.

Biographies

Marko Aho did his PhD in 2002 on cultural meaning of central Finnish early pop-music icons. In 2002 – 2006 he worked as a research director at the Dept. of Music Anthropology, University of Tampere, and in 2007 – 2011 as the director of Folk Music Institute. He has published papers in journals such as Popular Music, Popular Musicology Online, Music Performance Research, and on national level, Yearbook of Ethnomusicology of the Finnish Ethnomusicological Society. He has been a member of the board of the Finnish Ethnomusicological Society for several years.

Tuomas Eerola is a Professor of musicology at the University of Jyväskylä, Finland. His PhD (2003) topic concerned with cross-cultural correlates of melodic expectancy. His research interest lies within the field of music cognition and music psychology, specifically in perception of emotions, rhythm and timbre. He approaches these topics by combining computational modelling and empirical experimentation. He has published in a variety of journals (e.g., Cognition, Music Perception, Musicae Scientiae, Psychology of Music, and PNAS). He has been a member of the board and the president of the Finnish Musicological Society for several years.