

Measuring *Aksak* Rhythm and Synchronization in Transylvanian Village Music by Using Motion Capture

FILIPPO BONINI BARALDI [1]

Ethnomusicology Institute (INET-md), FCSH, Universidade NOVA de Lisboa

EMMANUEL BIGAND

CNRS UMR 5022, Research Laboratory on Learning and Development, Université Bourgogne Franche-Comté

THIERRY POZZO [2]

INSERM-U1093, Université Bourgogne Franche-Comté

ABSTRACT: Techniques based on motion capture can be useful in analyzing a wide range of musical styles and practices: in this case, Transylvanian village music. We focused on a repertoire known as ‘Gypsy songs of sorrow’, played by professional Gypsy musicians during parties and celebrations of their own community. Two parameters were the object of study: rhythmic duration, and synchronization between musicians (a violinist and a viola player). Results show that rhythm is a local variant of *aksak* and is based on two duration units (S=short, L=long) which respect the formula $2:3 < S:L < 3:4$. Performances are characterized by large variations of the S:L ratio from period to period, which have an expressive function. Tracking the bow’s movements with motion capture techniques allowed to show that these variations are related to a swinging interpretation, which also involves a voluntary asynchrony between the two musicians.

Submitted 2014 December 15; accepted 2015 September 23.

KEYWORDS: *Ethnomusicology, gesture, motion capture, aksak, synchronization*

INTRODUCTION: AKSAK RHYTHM AND SYNCHRONIZATION IN TRANSYLVANIAN GYPSIES’ ‘SONGS OF SORROW’

IN order to always fulfill their clients’ musical requests, professional Gypsy musicians in Transylvania (Romania) are supposed to know an exceptionally large repertoire of ‘dance tunes’ (Romanian: *de joc*) and ‘listening tunes’ (*de ascultare*). Both genres are played in many different social contexts, such as weddings, funerals, baptisms, and spontaneous parties, in urban as well in rural areas. Listening tunes are also called ‘table songs’ (*de masă*), because they are played while the clients are sitting around the tables, in the communal house, singing, eating and drinking among friends (see Figure 1). A sub-group of table songs, which are usually played only for a Gypsy audience, are locally called ‘Gypsy songs of sorrow’ (*tiganesc de jale*). These tunes are usually played at Gypsy funerals, but also at the end of Gypsy wedding parties, when the sun rises, or during small parties in the Gypsy neighborhood. They often trigger tears from both musicians and listeners.[3]

Figure 2 presents a musical transcription of one of these Gypsy songs of sorrow, based on a CD recording published in 1998 (see discography) which can be heard here (Audio Ex1). As it is often the case for the instrumental repertoire of this region, this tune has no title. In this article, we will therefore refer to this recording with the abbreviation ‘CD’. The complex melodic figures played by the violinist, rich in ornaments, the chords played by the *contră* (a three-string viola), and the harmonic line played by the double bass may be observed.[4] The global structure of this tune is rather simple. In the first part (A, bars 1 and 2), the theme is played in the key of C (motif a’, bar 1), and is then transposed in the key of D (motif a”, bar 2). Part A is played twice, followed by part B (motifs x’, x”, y’, y”, z’ and z”, bars 5 to 10), where the melody develops by following a descending melodic line, and a final part C (motifs c’ and c”, bars 11

and 12) ends with a cadence on the new tonal center (key of A). Throughout the whole tune, the viola and the double bass play the harmonic accompaniment following the same rhythmical pattern. The presence of major chords only, which follow the descending profile of the melodic line, and the absence of any kind of additional chords (such as dominant 5th or dominant 7th), suggest that this represents an old rural melody, one that appeared prior to the enormous musical production of the Hungarian composers of the romantic period (Sárosi, 1978).

If the harmonic and melodic structure of these tunes are rather easy to detect and simple to analyze, three other musical parameters are rather unusual for a western ‘musical ear’. These are: a) the rhythmical system, belonging to the family of *aksak* and more generally of non-isochronous rhythms; b) the slight asynchrony between the melody and the harmonic accompaniment; and c) a particular way of elaborating the melodic line with ornaments and passing notes, locally called ‘sweetness’ (*dulceață*). We believe that the study of these musical parameters offers an important opportunity to show how these musicians develop elaborate musical skills and strategies of playing together that differ in subtle ways from those used in the western culture. Thus, the key aim of the present article is to contribute to an emerging non-ethnocentric analytical musicology, rooted on empirical methods (London, 2012). In this article, we will focus on the analysis of the first two parameters, the third being analyzed elsewhere (Bonini Baraldi, 2015).

Two studies were designed, both relying on empirical measures of tone durations and onset timings. The aim of Study 1 is to understand the rhythmic structure of the Gypsy songs of sorrow by precisely measuring the durations of the chords played by the viola. Statistical analysis is used here to calculate the mean values of the S:L ratio in 8 different songs of sorrow (S and L are the two basic rhythmic units that define *aksak* rhythms: S = Short, and L = Long). These measures should facilitate the comparison of the Transylvanian slow *aksak* with other similar rhythms observed in other regions of the world, thus enlightening the on-going theoretical debate on the *aksak* rhythmical system (e.g. Cler, 1994; 1998; Arom, 2004). The aim of Study 2 is to precisely measure the timing asynchronies among the violin part (melody) and the viola part (harmonic-rhythmical accompaniment), which produce an effect comparable to the idea of ‘swing’ or ‘groove’ in jazz. Our objective is to determine whether these asynchronies are related to the musicians’ expressive intentions, and how they affect the underlying *aksak* rhythmical system. This study should therefore contribute to current research on expressive deviations in ensemble performances (Keller, 2014; Wing *et al.*, 2014), focusing on inter-individual co-variations rather than on an isolated performer.



Fig. 1. Gypsy musicians playing ‘table songs’ (*cîntec de masă*) for their Hungarian clients during a banquet in the village communal house (Ceuaș, Transylvania, May 1st, 2004. Photo: Filippo Bonini Baraldi).

The image displays a musical score for a piece titled 'CD', which is a Gypsy song of sorrow. The score is presented in a three-staff format: the top staff is for violin, the middle staff is for *contră* (viola), and the bottom staff is for double bass. The tempo is marked as $\text{♩} = 75$. The score consists of six systems, each containing two measures. The first system begins with a treble clef and a key signature of one flat. The violin part features a melodic line with various ornaments, including wavy lines and double dots, and includes a triplet of eighth notes. The *contră* and double bass parts provide harmonic support with chords and single notes. The piece concludes with a double bar line at the end of the sixth system.

Fig. 2. Musical transcription of the first cycle of the Gypsy song of sorrow ‘CD’ (see discography and listen to the track Audio Ex1). Top line: violin; middle line: *contră* (viola); bottom line: double bass.

STUDY 1: MEASURES OF AKSAK RHYTHM

Hearing these slow listening tunes, it sounds as though the rhythmical system belongs to the family of *aksak* (from the Turkish, meaning ‘limping’). Following Brăiloiu (1973), a rhythm is an *aksak* when: (1) it is periodic and (2) each period is made up of the combination of two categorically distinct duration units (Short = S and Long = L), nominally in a ratio of S:L = 2:3. *Aksak* rhythms are usually indicated by the series constituting the period, and by taking the value of 2 as duration for S and 3 for L; for example 2.2.3, 2.3.2, 2.2.2.3, etc. In reference to the *aksak* typologies proposed by Brăiloiu (1973), the rhythmical system of the Gypsy songs of sorrow would belong to his type IIa6:



Throughout this article, we will use the expression ‘rhythmic units’ to indicate both S and L, and the term ‘period’ to indicate the repeated sequence given by S.L. In our musical transcription (see Figure 2), one measure includes therefore two periods, that is, four rhythmic units which comprise the sequence S.L.S.L. Following Brăiloiu (1973), the rhythmic unit S is notated as a quarter note, and L as a dotted quarter note. It can be observed from the transcription that the basic rhythmical pattern of the viola and the double bass relies on quarter notes (S) and dotted quarter notes (L), with a single tone corresponding to each rhythmic unit. In some cases, the L rhythmic unit is decomposed in an eighth note and a quarter note (measures 5 and 9). In some other cases, the violist introduces eighth notes and sixteenth notes that should be considered as ornaments (measures 2, 6 and 10). Finally, in this recording (CD), the quarter note (S) is played at a tempo of about 61 bpm (IOI = 984 ms), and this tempo remains approximately stable throughout the whole tune.[5] We are therefore in the presence of a slow *aksak*, less studied than the fast *aksak* rhythms, which are very common in the Southern Balkan regions, in Turkey and in Asia Minor.[6]

Aksak rhythms have been at the center of controversial debates in French ethnomusicology, and there is no agreed upon theoretical framework to describe them (see Arom, 1992; 2004; Cler, 1994; 1998; Bouët, 1997; Cler & Estival, 1997). The main question concerns the way we should understand *aksak*: in terms of two distinct beat classes, with an inherent ratio of approximately 2:3, or of an underlying pulse, which means that musicians rely on one time unit only (an eighth note), S and L being multiples of this unit. In this second case, the *aksak* pattern 2.3.2.3 could be transcribed in 10/8, and the pattern 2.3 in 5/8. A related question concerns the way of understanding the ‘elasticity’ of *aksak*, i.e. the large variations in S:L proportions that seem to characterize *aksak* rhythms, especially in Romania (Bouët, 1997; Haplea, 2005).

Precise measurement of the S and L durations and of the S:L ratio in live performances may offer an important indicator of how an *aksak* rhythm is locally conceived by musicians. This may help to compare *aksak* in different regions of the world and to advance stronger theoretical models. Yet, only a few studies have dealt with the empirical measurement of S:L proportions (Cler, 1994; Cler & Estival, 1997; Polak, 2010; Polak & London, 2014; Jankowsky, 2013; Goldberg, 2014).

By analyzing his own recordings of Turkish music, Cler showed that the S:L ratio varies according to the tempo, the instruments played, and the performance context (Cler, 1994; 1998; Cler & Estival, 1997). In an article emblematically titled ‘*Aksak*, the catastrophes of a model’, Cler argues for the need to accept the idea of ‘a fragmented model, offering now coincidences with the ‘pure’ abstract form, now distortions, trans-formations – in the etymological meaning of the term’ (Cler & Estival, 1997, p. 75). The author puts forward a general *aksak* model, constructed *a posteriori*, with one or more distinct cognitive models intended to represent the way musicians locally conceive the rhythm.

Research by Bouët (1997) confirmed the ambiguity and the flexibility of the *aksak* model. Bouët transcribed the same *purtata*, a dancing tune of central Transylvania, by using the proportions of S:L = 2:3 (which he calls ‘orthodox *aksak*’), and by using the proportions of S:L = 3:4 (which he calls ‘heterodox *aksak*’). By playing these musical transcriptions with a standard sound software (Midi Composer), the author affirms that ‘we cannot perceive a difference that would give more validity to one of the two’ (Bouët, 1997, p. 119). Bouët did not run perceptual tests in order to determine if Romanian listeners could distinguish one version from the other, and this conclusion is based only on his personal familiarity with Transylvanian rural music. Even if the author did not perform timing analyses to this *purtata*, he argues that different S:L proportions coexist and alternate in a same tune, and claims therefore that we cannot have a strict and unique conception of the *aksak* model.

The problem of *aksak* can be related to current research on non-isochronous rhythms (Temperley, 2000; Polak, 2010; London, 2012; Polak & London, 2014).[7] Non-isochronous rhythms have been documented by ethnomusicologists in many different regions of the world (Pantaleoni, 1972; Kolinski, 1973; Hood, 1982; Stone, 1985; Kubik, 1988; Clayton, 2000; Arom, 2004), but have been excluded from the western musical theory for a long time. More recently, empirical methods have been used to measure non-isochronous rhythms in non-western repertoires. Analyzing audio and video recordings of Mande drummers (Mali), Polak & London (2014) have shown that the rhythmical accompaniment in both *Ngòn* and *Bire* tunes involves a non-isochronous ‘Long-Short’ subdivision, with an average timing ratio of L:S:S = 41:31:28 and L:S = 59:41 respectively. This non-isochronous subdivision seem to be largely independent from the performance context, the tempo, and the musicians’ expertise. The authors conclude that even if Mande drumming is based on a isochronous beat, a strongly non-isochronous subpulse (L.S.S or L.S) subdivides the isochronous beat. Conversely, Kvifte (2007) and Johansson (2009), who analyzed rural dance music from northern Europe, showed that both metric beat and subdivision can be non-isochronous. This hypothesis is also demonstrated by Jankowsky (2013), who measured timing patterns in Tunisian healing ceremonies and showed that the rhythmical patterns are not reducible to an underlying isochronous common fast pulse, a pulse which would serve as the lowest metric referent level.

Starting from this basis, Study 1 deals with the precise measurement of the *aksak* S:L proportions in eight different Gypsy songs of sorrow. In relation to the empirical studies on *aksak* and non-isochronous rhythms previously cited, which are based on timing measurements of the audio signal, the current study combines audio analyses (part 1) and motion capture based timing analyses (part 2).

Study 1, Part 1: Rhythmical Measures Using Sound Analysis

METHOD

In July 2007, two Gypsy musicians from the Transylvanian village of Ceuaş, a violinist (Sanyi) and a viola player (Csángáló), were invited to give a workshop at the Cité de la musique in Paris, France. On this occasion, we organized a recording session in the fully equipped sound studio of the Cité de la musique. The viola player who took part in this recording session was the same musician who recorded this tune in 1998 (CD), while the violinist was a different musician. They are among the best performers of Transylvanian village music, they were born in 1959 and 1951, and have played together since they were children. Among many other tunes, we recorded the same Gypsy song of sorrow published in the 1998 CD and transcribed in Figure 2. We will refer to this newer recording, which can be listened here (Audio Ex2), with the abbreviation ‘CM’ (standing for Cité de la musique).

The two musicians were placed in two separate rooms in order to obtain stereo recordings with the two instruments in separate channels (violin left, viola right). The musicians could see each other through a glass window and hear each other through headphones. This procedure was used in order to facilitate the measurement of the rhythmical pattern, since in the viola part each chord corresponds generally to one of the two rhythmic units S and L (see Figure 2). Timing durations of S and L could therefore be obtained from the viola waveform signal alone (Audacity ver. 2.0.6).

DATA CODING

Recording the viola part in a separate audio channel did not solve the well-known problem of onset detection for string instruments, that is, the problem of how to choose the points in the audio signal that best correspond to the beginning of each S and L rhythmic unit. Moreover, in this style of playing, the technique of the viola is based on a continuous sound (the bow never leaves the strings), and the perceived starting point of each rhythmic unit (which the musician plays as an accent) comes slightly after the change of direction of the bow.

We considered therefore three possible choices for determining the starting and ending point of each S and L rhythmic unit. A first possibility consisted in choosing the points of highest amplitude in the viola waveform signal, which would correspond to the accents played by the musician at the beginning of each new rhythmic unit. This measure is presumably consistent with the way the musician ‘thinks’ the rhythm, but it is hard to obtain with high precision. The second possibility consisted in choosing the lowest amplitude points, which corresponds to the moment in which the bow changes direction. This measure is easy to obtain with high precision, but does not correspond to the exact moment in which the beginning of

a new rhythmic unit is perceived. The third possibility was to choose the starting and ending points for S and L rhythmic units on the basis of perception, that is, by manually placing time markers on the waveform signal when the beginning of a new unit is perceived. This last method is much like the ‘P-center’ problem in psychological studies of vowel timing in language (e.g., Patel, L fquist, & Naito, 1999).

It is worth to note that the problem of choosing the correct starting points for S and L is partially overcome by the fact that we are interested here in calculating ratio and mean values. Choosing the points in a consistent way throughout the song would therefore give a reliable indication of the S:L ratios for each period and S:L mean values for the entire song. Yet, in order to assure a high degree of precision, we performed all three types of measures and we calculated the average of the three. In this article, we will refer to these three different types of measures with the expressions ‘Peak Amplitude Onset (PAO)’, ‘Minimum Amplitude Onset (MAO)’ and ‘Perceptual Onset (PO)’ (see Figure 3). All measures were obtained by a single researcher (Bonini Baraldi) who in the third case (perceptual measures) relied on his familiarity with the local musical style. In all three cases, we assumed that the ending point of one rhythmic unit corresponded to the beginning of the next rhythmic unit. Time markers have been placed on the waveform, and then the running time code was extracted by using the Audacity ‘extract markers time codes’ option. A simple Excel function has then been used to calculate the time durations of S and L (in seconds), the S:L ratio for each period, and the mean S:L value over 44 periods, which corresponds to the entire tune.

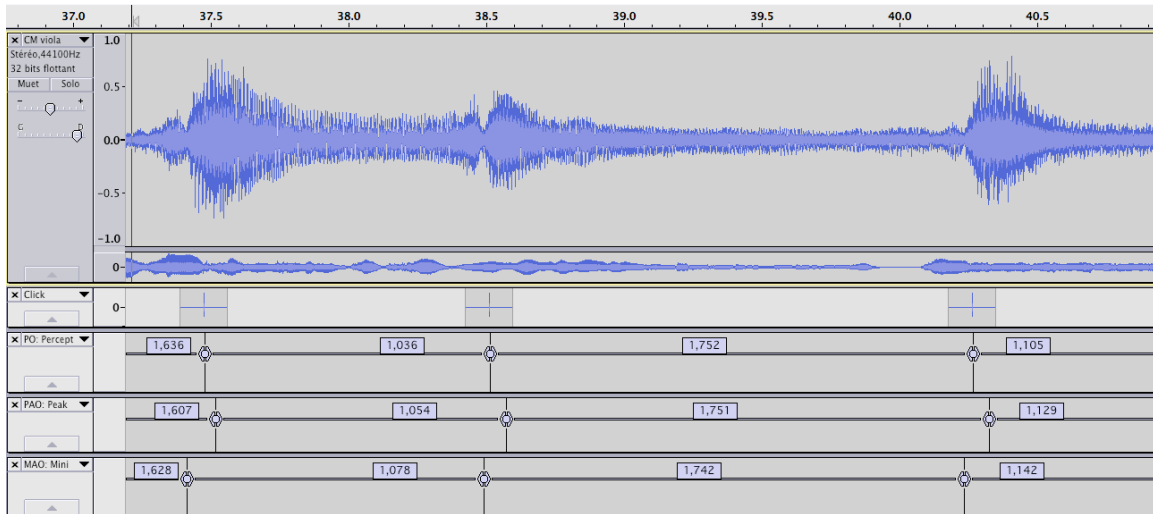


Fig. 3. Audacity (ver. 2.0.6) window. The upper part shows the viola wave signal of the Gypsy song of sorrow ‘CM’ (Audio Ex2). The bottom part shows three different possibilities for determining the starting points of the rhythmical units S and L, named respectively ‘Perceptual Onset (PO)’, ‘Peak Amplitude Onset (PAO)’, and ‘Minimum Amplitude Onset’ (MAO). PO measures of S and L durations were obtained by (manually) placing the time markers on the sound signal when the beginning of a new unit is perceived. PAO measures were obtained by placing markers in correspondence to the points of maximum amplitude of the sound signal, which correspond to the accent played by the musician at the beginning of each new rhythmical unit. MAO measures were obtained by placing markers in correspondence to the lowest amplitude points, which correspond to the moment in which the bow changes in movement direction.

RESULTS

Table 1 summarizes the mean values of S, L and S:L ratio for the Gypsy song of sorrow recorded at the Cit  de la musique (CM), obtained with the three measuring methods above described (PAO, MAO and PO). The average of the three sets of measures is also included in the table (in bold). Results show that S and L durations may vary largely throughout the performance (minimum S = .930 s at period 31 and maximum S = 1.186 s at period 5; minimum L = 1.461 s at period 43 and maximum L = 1.795 s at period 8). These large variations of S and L timings are not related to ‘global’ *accelerando* or *rallentando* patterns;

rather, they should be conceived as ‘local’ (at the level of each period) stretchings or shortenings of the S and L mean durations. These variations of S and L durations are in some cases linked to the rhythmical decompositions played by the viola player, such as in period 5, when the L rhythmical unit (dotted quarter note) is decomposed in one eighth note followed by a quarter note.

Table 1 also shows the S:L ratios, calculated for each period by dividing each S value by the corresponding L value. Firstly, results show that the S:L mean ratios in the three different sets of measures are very close one to each other, which means that we can consider each one as reliable. Secondly, the grand average S:L ratio for the entire song (44 periods) is $S:L = .656$ (st. dev. = .0435). This value is very close to the ‘orthodox’ *aksak* proportion ($S:L = 2:3 = .667$), which confirms that the rhythm of the slow Gypsy songs of sorrow may be considered as an *aksak*. Finally, the S:L ratio varies throughout the performance, ranging from a minimum value of $S:L = .539$ (period 8) to a maximum value of $S:L = .759$ (period 1). This means that the musicians take the liberty of changing the S:L proportion in the course of a performance. This variability may be due to structural factors (such as lengthening notes at the end of phrases), random variations, and errors, and is characteristic of most music performances, both for standard meters as well as non-isochronous meters such as *aksak*. Nevertheless, measures on one single tune (CM) do not allow us to understand if these variations in S and L durations and in S:L ratios are linked to specific passages of the music. In order to generalize these results, a larger set of measures was needed. This was obtained by using a different methodology: motion capture analysis of musicians’ gestures.

		<i>Peak Amplitude</i>	<i>Minimum Amplitude</i>	<i>Perceptual</i>	
		<i>Onset measures</i>	<i>Onset measures</i>	<i>Onset measures</i>	<i>Average</i>
		<i>(PAO)</i>	<i>(MAO)</i>	<i>(PO)</i>	
<i>S (Short) duration</i> <i>in seconds</i>	<i>MEAN</i>	1.056	1.082	1.060	1.066
	<i>ST. DEV.</i>	.0646	.0677	.0651	.0637
<i>L (Long) duration</i> <i>in seconds</i>	<i>MEAN</i>	1.640	1.612	1.636	1.629
	<i>ST. DEV.</i>	.0773	.0889	.0771	.0773
<i>S:L Ratio</i>	<i>MEAN</i>	.645	.673	.649	.656
	<i>ST. DEV.</i>	.0432	.0518	.0438	.0435

Table 1. Mean values of S, L and S:L ratio calculated on 44 periods of the Gypsy song of sorrow (CM) with three different measurement methods: ‘Peak Amplitude Onset’ (PAO), ‘Minimum Amplitude Onset’ (MAO) and ‘Perceptual Onset’ (PO). The average of the three measures is indicated in bold.

Study 1, Part 2: Rhythmical Measures Using Motion Capture

METHOD

In 2007, one week after the recordings at the Cité de la musique, the same two Gypsy musicians were invited to the Université de Bourgogne (Dijon, France) for a recording session that included motion capture technologies. Among other scientific tasks that motivated the use of these technologies, we believed that the 3-dimensional images of the musical gesture could be particularly useful for measuring the *aksak* rhythmical patterns of the Gypsy songs of sorrow. This is because, in the viola part, each bow stroke corresponds to one of the two rhythmic units S and L, the short duration being played with a downward bow stroke, and the long duration with an upward bow stroke. Following the movement of the marker placed at the top of the viola musician’s bow should therefore allow to measure the S and L durations for each period, and to determine the S:L ratio with a higher precision than the measures obtained from the audio signal (part 1).

Musicians were asked to play 30 s fragments (corresponding to 12 or 13 periods) of eight tunes issued from the local repertoire of Gypsy songs of sorrow. One tune, the same as CD and CM (see Figure 2), was played 3 times (Duo12, Duo13, and Duo14). In order to understand the influence of musicians’

expressive intentions on the timing durations of S and L, we asked musicians to play Duo12 and Duo13 ‘without sweetness’ (*fără dulceață*) and Duo14 ‘normally’, that is, ‘with sweetness’ (*cu dulceață*). The other tunes, named Duo18, Duo19, Duo20, Duo21 and Duo22, were all different. All tunes were played approximately at the same tempo (average: S at 55 bpm, IOI = 1091 ms).

The session lasted about 3 hours and the musicians, who are used to much longer musical performances, felt comfortable in the laboratory setting. The sound was recorded separately and the session was filmed with two additional standard cameras. Synchronization between sound and image was obtained by using a clapper board equipped with two additional retro-reflective markers. All audio recordings can be listened here (Duo12 = Audio Ex3, Duo13 = Audio Ex4, Duo14 = Audio Ex5, Duo18 = Audio Ex6, Duo19 = Audio Ex7, Duo20 = Audio Ex8, Duo21 = Audio Ex9, Duo22 = Audio Ex10), and a video of the laboratory session can be screened by running with a common internet browser the page ‘cinématiques’ of the interactive animation ‘Jouer la *jale*’ (performing sadness) available here (Animation Ex1).

DATA CODING

The two musicians were standing at the center of a circular region of 2,5 m radius, surrounded by the six infrared-emitting cameras attached to six tripods, 2 m from the ground on each side of the subject, at a distance of 3 m from each subject’s body. The movements of 38 retro-reflective markers (15 mm in diameter), placed at various anatomical locations on the body, were measured using an optoelectronic device, Smart (BTS, Milan, Italy). Kinematic parameters in three dimensions (X, Y, and Z) were calculated from successive frames taken at 10 ms intervals. Kinematic variables were low-pass filtered using a digital second-order Butterworth filter at a cutoff frequency of 5 Hz.

This computation allows obtaining a mathematical model of the markers’ trajectories in space, which can be viewed in 3 dimensions with the software ‘SMART Viewer’. Figure 4 shows the trajectory (in 2 dimensions) of a single marker, the one positioned at the top of the violist bow, during the performance of Duo14. The horizontal dimension of the plot represents time (in seconds, resolution 7 ms), and the vertical dimension represents the position of the marker in space (meters from the ground, resolution 2 mm).[8] The periodical cycle of the *aksak* rhythm is clearly recognizable. The peaks indicate a change in the bow direction, corresponding to the onset of a S rhythmic unit (superior peaks) and of a L rhythmic unit (inferior peaks). The webpage named ‘rhythm’ of the interactive animation ‘Jouer la *jale*’ (Animation Ex1) shows how this plot was obtained with the software ‘SMART Viewer’ (see Figure 4).

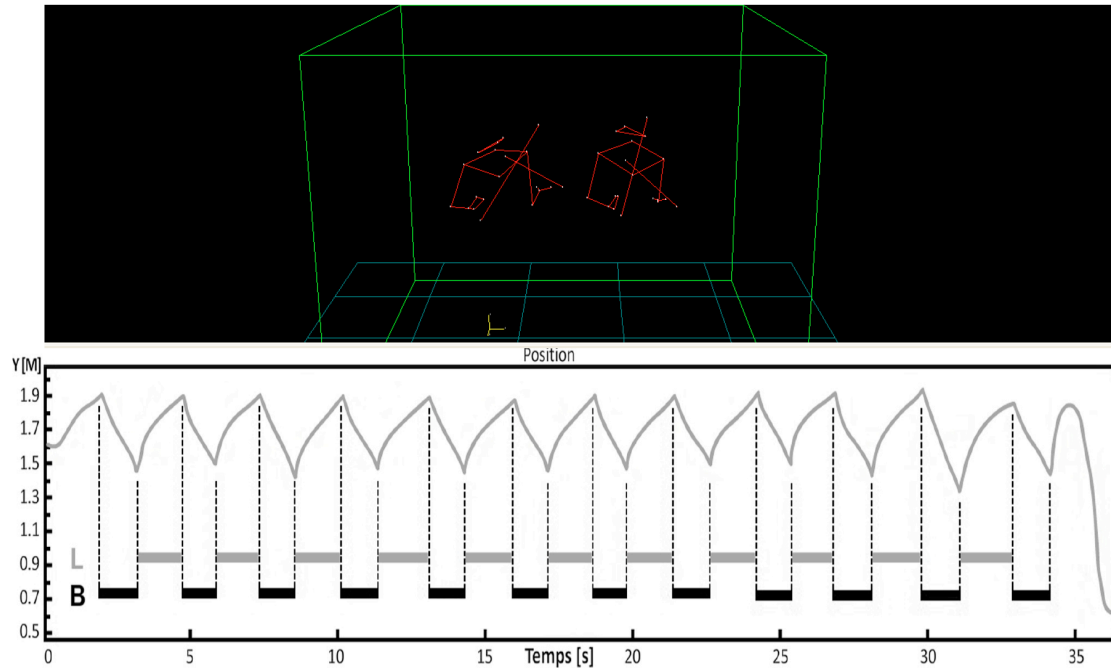


Fig. 4. ‘SMART Viewer’ window. Top: the musicians’ virtual silhouettes (violinist left, violist right). Bottom: rhythmical pattern of Duo14, (Audio Ex5, Animation Ex1) obtained by tracking the marker positioned at the top of the viola bow on the Y dimension. The horizontal dimension of the plot represents time (in seconds, resolution 7 ms), and the vertical dimension represents the position of the marker in space (meters from the ground, resolution 2 mm). ‘L’ stands for ‘Long’ rhythmic unit, and ‘B’ (abbreviation of the French ‘Brève’) for ‘Short’ rhythmic unit.

The two-dimensional representation of the upward and downward motion of the bow allows one to obtain the timing durations of S and L, by measuring the time interval between one peak and the next. It could be argued that this method does not take into account the fact that musicians’ body movements may introduce irregularities in the trajectory of the bow, which in turns would make difficult to precisely measure the timing durations of S and L rhythmic units. For example, if the musician bends his trunk while playing, the bow trajectory would change, and this would not allow a precise measure in time domain. Two arguments can be advanced to support the validity of this method. Firstly, in the local technique, musicians play in a very fixed position and upright posture, and expressive body movements are explicitly avoided. This is because when they perform in a professional context, professional Gypsy musicians are supposed to play for their clients (and as researchers, we were also considered as ‘clients’) without expressing their own feelings and emotions through body movements or facial expression (Bonini Baraldi, 2013). Figure 4 supports this ethnographic observation: the plot of the bow’s movements shows a very regular pattern, which confirms that the musician maintains a fixed position in the space (an exception can be observed at the last period of Duo14). Secondly, we are interested in ‘local’ measures of the *aksak* proportions (that is, at each period) and in S:L mean values for the entire song. Thus, we can assume that, for our sake, small changes of the musicians’ position are insignificant.

Nevertheless, in order to be sure that the viola bow’s movement give a precise indication of S and L timing durations, measures obtained for Duo14 with motion capture technologies (MC) were compared with those obtained on the same recording with the ‘Perceptual Onset’ method (PO) (Table 2).[9] The mean S, L and S:L ratio, calculated over 12 periods are very similar, which suggests that the MC measures can be taken as a reliable alternative method to sound analysis.[10]

		<i>Perceptual Onset measures (PO)</i>	<i>Motion capture measures (MC)</i>
<i>S (Short) duration in seconds</i>	<i>MEAN</i>	1.143	1.155
	<i>ST. DEV.</i>	.0441	.0665
<i>L (Long) duration in seconds</i>	<i>MEAN</i>	1.665	1.652
	<i>ST. DEV.</i>	.1139	.0955
<i>S:L Ratio</i>	<i>MEAN</i>	.688	.703
	<i>ST. DEV.</i>	.0389	.0463

Table 2. Comparison of timing measures for Duo14 (mean values), obtained with motion capture (MC) and with ‘Perceptual Onset’ (PO) methods.

RESULTS

Measures of S and L durations in Duo14, obtained with motion capture techniques, confirmed two results obtained in part 1 of the current study, where we analyzed the audio signal of a different recording of the same tune: the S:L average proportions (.703) are close to the *aksak* ratio (.667), and the durations of S and L rhythmic units may be stretched or shortened during the performance (for Duo14, $1.05 \text{ s} < S < 1.267 \text{ s}$; $1.541 \text{ s} < L < 1.758 \text{ s}$; and $.657 < S:L < .817$).

Timing measures based on motion capture (MC) were then conducted on the entire set of Gypsy songs of sorrow (8 tunes, see Table 3). The mean values of S:L ratio for the three performances of the same tune (Duo12, Duo13 and Duo14) are very similar (respectively .695; .696; and .703). An ANOVA performed with the 8 pieces as the within-subject variable and the S:L values as the dependent measure, revealed a main effect of piece $F(7, 77) = 3.05$, $p < .01$, caused by two pieces (Duo18 and Duo22) which have significant higher S:L values than all other pieces. Post hoc contrasts for the remaining duos (Duo12, Duo13, Duo14, Duo19, Duo20, Duo21) did not reach statistical significance. This provided some evidence that S:L values were highly similar for all duos, except the two mentioned above. In addition, all S:L mean values respected the formula $2:3 < S:L < 3:4$.

<i>Motion capture recordings</i>	<i>S:L mean values on 12 periods</i>	<i>St. dev.</i>
<i>Duo12</i>	.695	.0272
<i>Duo13</i>	.696	.0354
<i>Duo14</i>	.703	.0436
<i>Duo18</i>	.735	.0286
<i>Duo19</i>	.679	.0663
<i>Duo20</i>	.689	.0491
<i>Duo21</i>	.687	.0358
<i>Duo22</i>	.729	.0384
<i>Average on 8 tunes</i>	.702	.0203

Table 3. S:L mean values for eight Gypsy songs of sorrow, obtained with motion capture methods. Duo12, Duo13, and Duo14 are different takes of the same tune, played with different expressive intentions (Duo12 and Duo13 ‘without sweetness’ and Duo14 ‘with sweetness’).

Discussion on *Aksak* Measures

We argued that precise measures of the S:L proportions in live performances may help to understand how musicians conceive of *aksak* in different parts of the world, and therefore to advance more general hypothesis on this particular rhythmical system. While measures of *aksak* proportions have been obtained for Turkish village music (Cler & Estival, 1997) and for Bulgarian and Serbian music (Goldberg, 2014), to our knowledge Transylvanian *aksak* repertoire has only been analyzed qualitatively (Brăiloiu, 1973; Bouët, 1997; Haplea, 2005).

The main result obtained in Study 1 is that the rhythm of the Gypsy songs of sorrow is neither an ‘orthodox *aksak*’ (S:L = 2:3) nor an ‘heterodox *aksak*’ (S:L = 3:4, see Bouët, 1997), but rather an *aksak* which is ‘in between the two’, since it is characterized by the formula $2:3 < S:L < 3:4$. In other words, if we represent the Short rhythmic unit (S) with a quarter note, as in the *aksak* convention, the average duration of the Long (L) is a little shorter than a dotted quarter note. The fact that this formula applies to all eight songs of sorrow analyzed here, may suggest that the musicians conceive two ‘blocks’ of durations (S and L), which are largely independent one from the other, that is, they do not rely on a common underlying pulse. In other words, the *aksak* rhythm – or at least, this Transylvanian version of *aksak* – should be regarded in terms of two independent duration units, as proposed by Brăiloiu (1973), rather than in terms of a unique smaller subpulse, as suggested by other scholars such as Arom (2004). Nevertheless, comparison with other studies should be advanced carefully, since we are in presence here of a slow tempo *aksak*, less studied than the fast tempo *aksak* rhythms analyzed by these authors.

Ethnographic observations seem to support the hypothesis of the co-presence of two different beat classes. Firstly, musicians and listeners tap the beat by following this ‘limping pulse’ (i.e. they tap the S and the L at each period), and not by breaking up S and L into a smaller pulse. Secondly, in some Transylvanian villages, these slow songs may be danced. In this case, the dancing steps clearly follow the pattern S.L, one foot marking the beginning of the S rhythmic unit, the other foot marking the beginning of the L rhythmic unit. Even if tapping behavior is not a transparent measure of timing mechanisms, this suggests that the Transylvanian slow *aksak* is not ‘bodily’ decomposed in smaller pulses. Thirdly, the violist plays eights notes very rarely, and usually each bow stroke corresponds to one of the two basic rhythmic unit (S being played with a downward bow stroke, and L with an upward bow stroke).[11]

The second main result that emerged from this study concerns the variability of the S and L timing durations and of their ratio (S:L) throughout the entire performance of a same tune. This result differs from the one obtained by Cler & Estival (1997), who, in the case of Turkish *aksak*, observed a stability of S:L proportions in a single tune. The fact that Gypsy songs of sorrow are played in a slow tempo, in opposition to the Turkish fast dance tunes analyzed by Cler & Estival (1997), may leave more freedom to the performers in introducing timing deviations. We believe that in the Gypsy songs of sorrow, these deviations – that is, the liberty to smoothing out or lengthening the duration of both S and L rhythmic units – may be linked to an expressive function. Indeed, as the local expression used to name this repertoire suggests (*cîntec de jale*, i.e. songs of sorrow), the primary social function of these slow songs is to elicit listeners’ emotions, and eventually to make them cry. But if we accept that the slow *aksak* used in songs of sorrow is subject to ‘expressive timing’ (London, 2012), one question can be raised: how do musicians jointly produce these variations in timing? Measures obtained in study 1 show that musicians, in this particular repertoire, do not follow global patterns of *rubato*, *accelerando* or *rallentando* that in other musical traditions, such as the western one, are commonly associated to an expressive function.[12] Would therefore Gypsy musicians follow other forms of expressive timing, relying on different types of temporal coordination among parts? Our hypothesis is that the timing deviations observed in study 1 are the effect of a ‘swinging’ way of playing, intended as a voluntary temporal a-synchronization between the two performers. Motion capture technologies, associated to musicological analysis, were used in the following study in order to verify this hypothesis.

STUDY 2: SYNCHRONIZATION BETWEEN MELODY AND ACCOMPANIMENT

By lending an attentive ear to the Gypsy song of sorrow transcribed in Figure 2 (Audio Ex1), it sounds as though the melody (violin) is slightly out-of-synch with the harmonic-rhythmical accompaniment (viola and double bass), that is, some notes of the melody either anticipate or follow the corresponding chords.

The aim of study 2 is to understand the ‘nature’ of these asynchronies, and particularly to determine if musicians introduce these asynchronies randomly, or rather according to some musical rules (such as cadences, section passages, etc.).

The problem of synchronization in time of different parts of a musical ensemble has recently raised the interest of musicologists and music psychologists, especially those focusing on musicians’ ‘expressive deviations’. If it has been widely demonstrated that musicians introduce variations to the nominal values (such as those indicated in the musical score) of parameters such as tempo, intensity, articulation and sound quality for expressive purposes (among many others, see Gabriëlsson, 1995; Bonini Baraldi, Rodà, & De Poli, 2006), an additional aim of this emerging research area is to understand expressive variations in terms of inter-individual co-variation. In other words, this new research task consists in understanding how ‘experienced ensemble musicians coordinate their actions to bring about the optimal co-variation of expressive performance parameters’ (Keller, 2014; p. 260).

In western classical music, the basic assumption is that this ‘optimal co-variation’ tends toward a perfect synchrony of the musical parts. For example, if the leading violin of a string quartet is lengthening or shortening a tone duration, the other musicians are supposed to apply the same variations in order to keep the parts in synchrony. An empirical study by Wing *et al.* (2014) has focused on how musicians of a string quartet achieve synchrony while playing together, and has proposed a model that aims to explain how they correct asynchronies that may emerge in live performances.

The way musicians conceive of time asynchronies – also called ‘vertical timing deviations’ in order to distinguish them from the timing deviations of a single part, named ‘horizontal timing deviations’ (Keller, 2014) – depends on many factors which include musical style and cultural context. In some cases, an effort is made to keep parts separate in time, rather than synchronized. This is the case of jazz and of some popular musics such as hip-hop and funk, where the effect commonly known as ‘swing’ or ‘groove’ seem to be related to a slight a-synchronization among parts, especially between bass and drums (see Benadon, 2006; Hove *et al.*, 2007; Butterfield, 2010). Keil (1987) proposed the expression ‘participatory discrepancies’ to name these time asynchronies, introduced voluntarily by musicians in order to obtain a ‘productive tension’, or ‘drive’ that generates a sense of swing or groove.

In most cases, these asynchronies are introduced in a systematic way, that is, one part is constantly ahead of the other parts, or conversely, after the other part. The first case, known as the ‘melody leads phenomenon’, has received considerable attention by scholars, mainly working on the western classical repertoire (Palmer, 1989; Repp, 1996; Goebel, 2001; Keller & Appel, 2010). The second case (the melody shortly behind the accompaniment) is probably more frequent in jazz, when a ‘laid back’ feeling is sought (Friberg & Sundström, 2002). In both cases, the systematic asynchronies among parts are believed to be an indication of leader-follower relations that operate in musical ensemble (e.g., the leader comes first; see Keller, 2014).

While in the previous cases small timing asynchronies are voluntarily introduced as an expressive means, in other cultural contexts a-synchronization may be much more ‘radical’ and may have a different function. Ethnomusicologists have described music performances in which people play or sing, in the same acoustic space, by voluntarily following different beats. These large-scale intra-group or inter-group asynchronies may be related to the need for affirming personal or group identity (Seeger, 1987; Rappoport, 1999; Lucas *et al.*, 2011), to the expression of religious devotion (Martinez, 1996), or to agonistic interplay (Martinez, 1992; see discography). French ethnomusicologists proposed the term ‘Polymusics’ in order to name the case in which different musics are voluntarily performed simultaneously without being temporally coordinated (Rappoport, 1999; Beaudet *et al.*, unpublished paper). It is worth to note that polymusic is not particularly exotic, since it is also common in western musical practices such as free-jazz, carnivals and techno parades. Future cross-cultural research may compare different cases of a-synchronized playing, and may ask if the tendency to converge toward synchrony – a phenomenon known as ‘entrainment’ – is a musical universal or rather a matter of culturally inflected listening habits.

Timing asynchronies among musical parts may therefore range from the micro level (a few milliseconds) to the macro level (groups who play completely a-synchronized). In both cases, different methodologies have been used in order to detect, to measure and to compute timing asynchronies. Analysis of the sound signal, either starting from real recordings or from controlled experiments, is by far the most common methodology used (Goebel & Palmer, 2009; Moore & Chen, 2010; Loehr & Palmer, 2011; Marchini *et al.*, 2012; Wing *et al.*, 2014). This method requires one to precisely detect the tone onsets in each part, especially when different instruments are playing together. In other cases, analyses of asynchronies have been conducted on the basis of video recordings (Lucas *et al.*, 2011). This method

presents the same inconvenience of audio analysis (determination of tones' onsets) as well as being difficult to compute. We believe that motion capture technologies may be particularly suitable for analyzing interpersonal coordination of musical ensembles. In the present study, these technologies were used to detect and measure timing asynchronies between the melody (violin) and the harmonic-rhythmical accompaniment (viola).

Measuring asynchronies by using motion capture

METHOD

Three different motion capture recordings of the same tune (the Gypsy song of sorrow transcribed in Figure 2) were used to detect and measure asynchronies. These takes (Duo12, Duo13 and Duo14, corresponding to Audio Ex3, Audio Ex4 and Audio Ex5) were recorded one after the other in the same session at Université de Bourgogne (see Study 1), and were played approximately at the same tempo (average: S at 55 bpm, IOI = 1091 ms). While in Duo14 no special performance indication was given to musicians, we asked them to play Duo12 and Duo13 'without sweetness' (*fără dulceață*) in order to compare tunes played with different expressive intentions. The viola accompaniment is in this case the same as for the other tunes, but the violin part in Duo12 and Duo13 is performed with fewer ornaments, fewer passing notes and less vibrato. Comparison of Duo12, Duo13 and Duo14 should therefore allow to observe if asynchronies vary due to the musical structure and the musicians' expressive intentions.

DATA CODING

The movements of the marker placed at the top of the violin bow have been now compared to the movements of the marker placed at the top of the viola bow (see Figure 5, top). These patterns were projected on the time axis in order to obtain a musical transcription of both the melody and the harmonic accompaniment (see Figure 5, bottom). The violin technique, which involves the use of distinct bow strokes for each single note, makes the task easier: almost every variation in the direction of the violin bow corresponds to the beginning of a new note. The interactive animation 'Jouer la jale' (see Animation Ex1, page 'swing') shows how this plot was obtained with the software 'SMART Viewer'.

The trajectories in space of the two bows allow to qualitatively observe the asynchronies between the melody and the accompaniment (see Figure 5, bottom). These were then precisely measured (in ms) by subtracting the onset times of the violin from those of the viola. For each asynchrony detected, we also calculated the percentage of its duration in relation to the corresponding rhythmic unit.

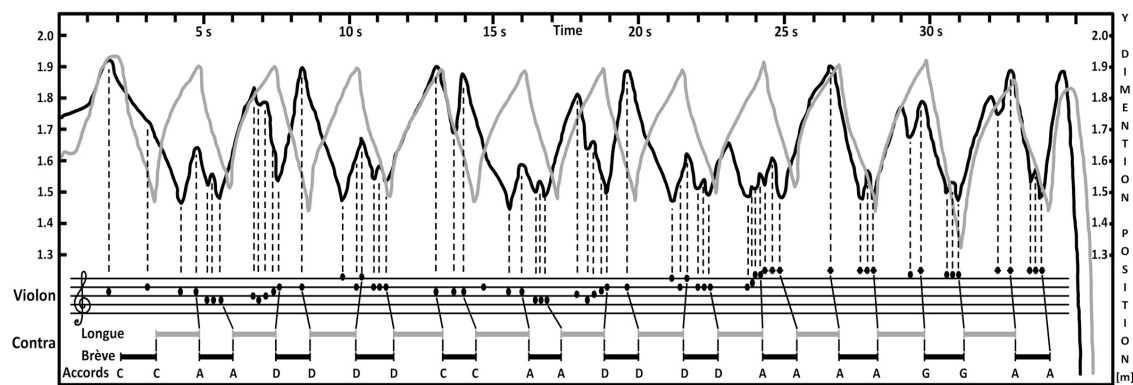


Fig. 5. Top: superposition of violin (black) and viola (gray) bow movements for the Gypsy song of sorrow Duo14 (Audio Ex5, Animation Ex1). The horizontal dimension of the plot represents time (in seconds, resolution 7 ms), and the vertical dimension represents the position of the marker in space (meters from the ground, resolution 2 mm). Bottom: transcription of melody (violin) and harmonic-rhythmical accompaniment (viola) obtained from changes in bow directions. The lines linking melody to accompaniment highlight the asynchronies between the two performers.

RESULTS

Tracking the movement of the violin and viola bows (see Figure 5) makes it easy to discern the different types of asynchronies existing between the two musicians: at almost each period, the chords anticipate or follow the corresponding tone of the melody. In other cases, such as at the beginning of period 9, the two musicians are perfectly synchronized. In this article, we will use the abbreviations ‘Posterior Asynchrony’ (PA) for the first case (viola after the violin), ‘Anterior Asynchrony’ (AA) in the second case (viola before the violin), ‘Synchrony’ (SY) in the third case (the two parts together), and ‘Non available’ (N) when it was not possible to determine the type of asynchrony. Since quantitative measures of asynchronies were obtained by subtracting the onset timing of the violin tone from the onset timing of the corresponding viola chord, PA are positive in sign, AA are negative, and SY are equal to 0.

Table 4 sums up the measures of asynchronies found in an entire cycle of Duo14, which is 12 periods long (corresponding to 24 rhythmic units) and is composed of the musical motifs $a' / a'' / a' / a'' / x' / x''$. Results show that the musicians introduce PA more frequently (17 occurrences) than AA (3 occurrences) and that they are rarely perfectly synchronized (1 SY). The same types of asynchronies are present in two different occurrences of the motifs a' and a'' , which suggest that asynchronies may be related to the musical structure of the song. PA vary from a minimum of 50 ms to a maximum of 550 ms. AA are introduced exclusively in correspondence of a Short rhythmic unit (i.e, at the beginning of a new period, such as in periods 3, 7 and 8), they are generally smaller than PA and their variation is lower. For both AA and PA, the percentage of asynchrony in relation to the beat duration ranges from 0% (perfect synchrony) to 51% (more than a half of the duration of the corresponding rhythmic unit). These measures suggest that two different types of asynchronies are present: ‘small-scale’ asynchronies (approximately 10% of the beat duration) and ‘large-scale’ asynchronies (approximately between 20% and 50%). Thus, musicians may conceive two different typologies of asynchronies, the first being very subtle, harder to perceive, and the second being very large and clearly audible even by a listener who is not familiar with the musical style.

Table 5 shows the amount of asynchrony (in ms) that was detected, for each period, in all three recordings of the same Gypsy song of sorrow (Duo12, Duo13 and Duo14), and Table 6 the percentage of asynchrony in relation to the duration of the corresponding rhythmic unit. Comparison among the different columns of these tables allows to observe if synchronization varies according to the musical structure and the musicians’ expressive intentions.

<i>Type of asynchrony</i>			<i>Amount of asynchrony</i>	
			In ms	In % of the beat duration
<i>Both S and L</i>	PA	<i>Mean</i>	225	19,6
		<i>St. dev.</i>	151	13,6
	AA	<i>Mean</i>	-136	11,8
		<i>St. dev.</i>	26	1,1
<i>S rhythmic unit</i>	PA	<i>Mean</i>	153	9,2
		<i>St. dev.</i>	72	4,9
	AA	<i>Mean</i>	-136	11,8
		<i>St. dev.</i>	26	1,1
<i>L rhythmic unit</i>	PA	<i>Mean</i>	285	25,4
		<i>St. dev.</i>	155	14,4
	AA	<i>Mean</i>	--	--
		<i>St. dev.</i>	--	--

Table 4. Types and measures of asynchronies (mean values) calculated over an entire cycle (12 periods) of the tune Duo14 (Audio Ex5). PA stands for ‘Posterior Asynchrony’ (chords follow the melody) and AA for ‘Anterior Asynchrony’ (chords anticipate the melody). Asynchronies were obtained (in ms) by subtracting the onset timings of the violin from those of the viola. For each asynchrony, was then calculated the percentage of its duration in relation to the corresponding rhythmic unit. No AA were detected in correspondence of L rhythmic units.

Musical Motif	Aksak Period	Rhythmic unit	Type and amount of asynchrony (in ms)		
			Duo12	Duo13	Duo14
a'	1	S	N	367	N
		L	N	N	N
	2	S	291	216	117
		L	533	492	375
a''	3	S	325	433	-159
		L	325	N	258
	4	S	-42	59	AA
		L	184	200	133
a'	5	S	267	241	166
		L	N	N	400
	6	S	250	0	125
		L	441	284	475
a''	7	S	442	400	-108
		L	392	N	258
	8	S	-133	58	-142
		L	141	142	266
x'	9	S	33	-58	0
		L	392	266	550
	10	S	333	367	292
		L	350	342	50
x''	11	S	375	225	91
		L	150	125	108
	12	S	217	250	125
		L	292	42	267
<i>Mean</i>			PA = 301.7 AA = -87.5	PA = 225.5 AA = -58	PA = 225.3 AA = -136.3
<i>St. dev.</i>			PA = 121.3 AA = 64.3	PA = 149.4 AA = 0	PA = 151 AA = 26

Table 5. Types and measures (in ms) of asynchronies detected, for each period, in three different recordings of the same Gypsy song of sorrow, played with different expressive intentions (Duo12 and Duo13 ‘without sweetness’, Duo14 ‘with sweetness’). PA stands for ‘Posterior Asynchrony’ (chords follow the melody), AA for ‘Anterior Asynchrony’ (chords anticipate the melody) and N for ‘data not available’.

Musical Motif	Aksak Period	Rhythmic unit	Type and amount of asynchrony (in % of the beat duration)		
			Duo12	Duo13	Duo14
a'	1	S	N	N	N
		L	N	N	N
	2	S	N	13.7	7.5
		L	52.4	46.1	35.7
a''	3	S	22.4	28.4	13.5
		L	30.2	N	22
	4	S	3.8	3.8	N
		L	16.6	19.4	11.6
a'	5	S	15.9	14.8	9.4
		L	N	N	35
	6	S	16.3	0	7.3
		L	40.7	25.6	41.9
a''	7	S	28.2	25.7	9.9
		L	34.3	N	23.6
	8	S	11.4	3.9	12.1
		L	12.1	13.3	22.6
x'	9	S	2	5.7	0
		L	38.2	26.2	51
	10	S	21.8	23.5	18.9
		L	30.4	29.5	4.1
x''	11	S	22.8	13.5	5.2
		L	13.3	11.3	8.8
	12	S	13.2	14.9	6.9
		L	27.2	3.6	22.4
<i>Mean</i>			22.7	17	17.6
<i>St. dev.</i>			12.7	11.63	13.6

Table 6. Types and measures of asynchronies (in percentage of the corresponding rhythmic unit) detected, for each period, in three different recordings of the same Gypsy song of sorrow, played with different expressive intentions (Duo12 and Duo13 ‘without sweetness’, Duo14 ‘with sweetness’). N stands for ‘data not available’.

(1) *Comparison of different occurrences of the same musical motif.*

While in Duo14 we found the same type of asynchronies in two different occurrences of the motifs a' and a'', in the other takes of the same melody this rule was not respected. Moreover, in the second occurrence of the motif a'' (periods 7 and 8), the musicians introduce very different asynchronies in almost every recording (Duo12: PA/PA/AA/PA; Duo13: PA/N/PA/PA; Duo14: AA/PA/AA/PA). This suggests that the musicians, while playing a tune, may synchronize differently even if the musical motif is the same.

(2) *Comparison of two consecutive takes of the same melody, played by the same musicians with the same expressive intention (Duo12 compared with Duo13).*

In Duo 12 there were 2 anterior asynchronies, 19 posterior asynchronies and 0 synchronies. The mean value of PA is 301.7 ms (22.7%), ranging from a minimum of 33 ms to a maximum of 533 ms (st. dev. = 121.3). As observed in the case of Duo14, posterior asynchronies are of two types: small-scale asynchronies (around 10% of the rhythmic unit duration), and large-scale asynchronies (from 20% to 50% of the rhythmic unit duration). Anterior asynchronies, which always occur at the beginning of a new period, are always ‘small’ (mean 121 ms, st. dev. = 64.4). In Duo 13, we detected 1 AA (58 ms, 5.8% of the rhythmic unit duration), 18 PA and 1 SY. In this case, mean PA is 225.5 ms (17.0%), ranging from a minimum of 42 ms and maximum of 492 ms (st. dev. = 149.5).

The comparison of Duo12 with Duo13 showed only four different types of asynchronies on a total of 22 rhythmic units (beginning of periods 4, 6, 8 and 9). In the other cases, the amount of asynchrony (in ms as well in percentage of the rhythmic unit duration) is rather similar, except in one case (period 10). Moreover, the general distribution of asynchronies along the song follows a similar pattern in Duo12 and in Duo13 (see Figure 6). These results suggest that musicians, when they play two consecutive takes of the same song, keep a ‘short-term’ memory of the way they synchronize, even if a certain amount of liberty is always present.

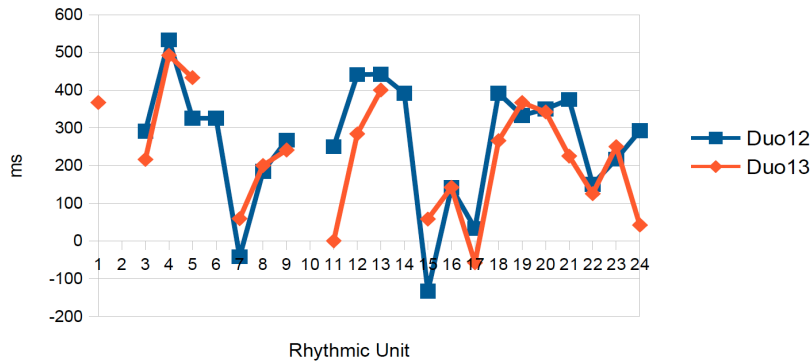


Fig. 6. Amount of asynchronies (in ms) between the viola and the violin in two performances of the same Gypsy song of sorrow played ‘without sweetness’ (Duo12 and Duo13). Since quantitative measures of asynchronies were obtained by subtracting the onset timing of the violin tone from the onset timing of the corresponding viola chord, Posterior Asynchronies (PA, the viola after the violin) are positive in sign, Anterior Asynchronies (AA) are negative, and Synchronies (SY) are equal to 0.

(3) *Comparison of two consecutive takes of the same melody, played by the same musicians with different expressive intentions* (Duo 12 and Duo13 compared with Duo14).

Further comparison was then made with Duo14, which is the same melody as Duo12 and Duo13 but performed with a different expressive quality (Duo14: ‘with sweetness’ vs. Duo12 and Duo13: ‘without sweetness’). The transcription presented in Figure 7 shows that the key difference in Duo14 stands in the elaboration of the basic melodic pattern by adding ornaments and passing notes.[13]



Fig. 7. Transcription of the Gypsy song of sorrow Duo14 (Audio Ex5), played ‘with sweetness’ (top), with indication of the chords played by the viola (middle) and the underlying melodic pattern (bottom).

As Figure 8 shows, large differences in the way musicians synchronize are present, especially at the beginning of period 3 and period 7. This means that the expressive intention locally called ‘sweetness’ (*dulceață*), which is strongly associated with a complex elaboration of the melodic part, freely introduced by the violinist, produces new types of asynchronies with the harmonic-rhythmical accompaniment. These differences can be explained by observing in detail the violin part in Duo14: the violinist introduces, at periods 3 and 7, four ornamental notes (notated as thirty-second notes) before the tone D of the melodic pattern, which are not played in Duo12 and Duo13 (see Figure 9). The introduction of this new musical element has the effect of delaying the onset time of tone D, which in Duo14 falls 159 ms after the chord D of the viola, and not before as in Duo12 (325 ms) and Duo13 (433 ms) (see Figure 10).

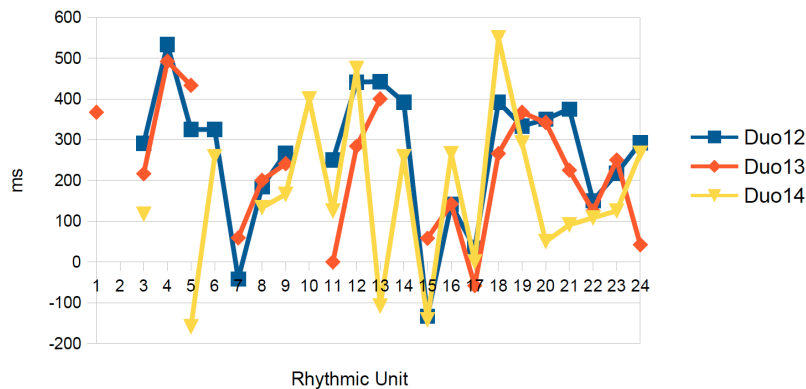


Fig. 8. Amount of asynchronies (in ms) between the viola and the violin in three performances of the same Gypsy song of sorrow played ‘without sweetness’ (Duo12 and Duo13) and ‘with sweetness’ (Duo14). Since quantitative measures of asynchronies were obtained by subtracting the onset timing of the violin tone from the onset timing of the corresponding viola chord, Posterior Asynchronies (PA, the viola after the violin) are positive in sign, Anterior Asynchronies (AA) are negative, and Synchronies (SY) are equal to 0.



Fig. 9. Detail of the violin part in period 3 of Duo 13 (top) and Duo14 (bottom). The fundamental difference in the version played ‘with sweetness’ (Duo14) stands in the elaboration of the basic melodic pattern by adding ornaments and passage notes. In this example, the violinist introduces in Duo14 four passage notes (notated as thirty-second notes) before the tone D of the melodic pattern, which are not played in Duo12 and Duo13.

To sum up, the number of differences in asynchronies increased from (1) to (3), that is, when two performances of the same musical motif were recorded with different expressive intentions. Besides a certain amount of rhythmic error or imprecision that may be introduced in any live performance, results suggest that musicians have a high level of freedom to introduce different types of asynchronies, even during the same performance. Thus, some regularity may also be observed, which means that asynchronies are not completely random. Indeed, in 9 occurrences, we found the same type of asynchrony for all five recordings of the same tune. In all 9 cases, the type of asynchrony is always a PA (the viola plays the chord slightly after the violin). This regularity concerns both S and L rhythmic units.[14]

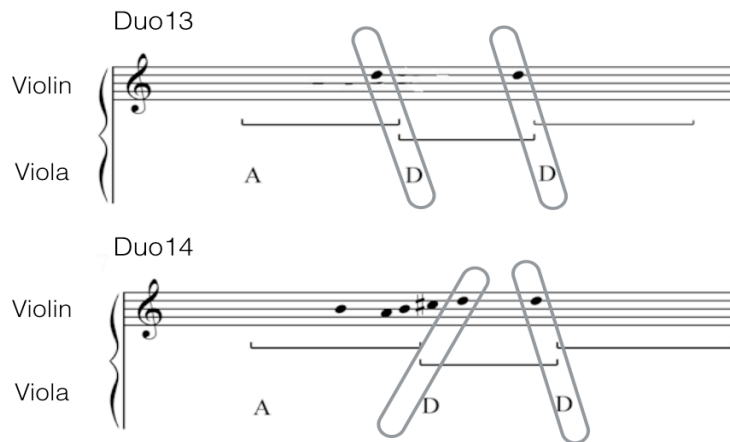


Fig. 10. The introduction of ornaments and passage notes in the version played ‘with sweetness’ (Duo14) has the effect of delaying the onset time of tone D of the melody, which in Duo14 falls 159 ms after the chord D of the viola (Anterior Asynchrony), and not before as in Duo12 (325 ms) and Duo13 (433 ms) (Posterior Asynchrony).

Discussion On Synchronization

How should we conceive the large presence of timing asynchronies in the slow Gypsy songs of sorrow, as well as their variation among different interpretations of the same tune? How do musicians themselves conceive this particular way of playing together? Is the time lag between the melody and the harmonic-

rhythmical accompaniment related to performers' errors, to their expressive intentions, or rather is it inherent to the musical structure of the song?

Before answering to these questions, it is important to note that, in the context of oral tradition, the concepts of 'nominal synchronous sounds', 'referent values' and 'deviations' – which are commonly used in current research on asynchrony in musical ensembles (Keller, 2014; Wing *et al.*, 2014) – do not apply in the same way as for written repertoires. In the former case, there are no written instructions (referent values) telling how parts should be coordinated one with each other, and we are often confronted with different interpretations of the same song which are all equally 'valid', provided they are accepted by the local community.[15] In the case of Transylvanian village music, there is no word corresponding to the idea of asynchrony, and this particular way of playing is largely implicit.[16] The results obtained with empirical methods should therefore be associated to ethnographic analysis, and comparison with research on other musical styles should be advanced carefully.

In the present study, what allowed us to assume that a chord 'should be' in correspondence with a particular tone of the melody, is the fact that the harmonic accompaniment, as it is played in Romanian rural areas, follows the tones played by the violinist (e.g., when the violinist plays a C, the violist play a major chord of C). Ethnographic field-work and author's familiarity with the practice of Gypsy music allowed us to determine the correspondences between the tones of the melody and the viola chords (see Figure 5). Relying on this 'emic' musicological analysis, asynchronies were intended and measured as timing deviations from these hypothetically synchronous sounds, even if there is no underlying fixed norm (such as a musical score), saying that these sounds 'should be' played together.

Starting from this basis, the type and the amount of asynchronies may be an indicator of how musicians conceive the coordination in time among musical parts. Results showed that, in the course of the performance of a Gypsy song of sorrow, asynchronies may range from a few milliseconds to more than half a second. We doubt that this variability is due to performers' errors, since Sanyi and Csángáló have been playing together for more than 40 years, and this Gypsy song of sorrow does not present for them any particular technical challenges. We supposed therefore that asynchronies should be understood in two different manners (large-scale and small-scale asynchronies), depending on their proportion with the corresponding rhythmic unit. Large-scale asynchronies (20% to 50% of the beat duration) may be highly related to the musical structure of the song, that is, to the way musicians conceive the melodic line in relation to the harmonic accompaniment. Indeed, comparison of two different takes of the same melody played with the same expressive intention (Duo12 and Duo13), showed that large-scale asynchronies are relatively stable among different interpretations of the same melody. This suggests that musicians assign a clear place to a tone of the melody in relation to the accompaniment, and that this place is often before the corresponding chord (predominance of posterior asynchronies). Conversely, small-scale asynchronies (less than 10% of the beat duration) may vary from one take to another, which suggests that these are less related to the musical structure, and more dependent from each particular interpretation.

It could be argued at this point that large-scale asynchronies may not be asynchronies at all, but rather rhythmic variations, such as when the violinist introduces a subdivision of the beat. In this case, we would need to set a threshold to distinguish rhythmic variations from asynchronies, for example, at about 15% of the beat duration. Nevertheless, our comparison of the two takes played 'without sweetness' with the one played 'with sweetness', shows that some rhythmic variations in Duo12 and Duo13 may become asynchronies in Duo14 (e.g. in periods 3, 4, 5, 7 and 10) and vice versa, asynchronies in Duo12 and Duo13 may become rhythmic variations in Duo14 (periods 8 and 12). Indeed, when the melody is played with more ornaments and passage notes (Duo14), the onset timings of the melodic line are altered, producing different types and amounts of asynchronies with the harmonic-rhythmical accompaniment. Thus, we believe that the terminology proposed here ('large-scale' vs. 'small-scale' asynchronies) better illustrates the fact that rhythmic variations may also be intended by musicians as (large-scale) asynchronies. This ambiguity in the placement of the melody in relation to the rhythmical pattern played by the accompanist is precisely what characterizes this oral tradition repertoire. In any case, both small-scale asynchronies and large-scale asynchronies seem to be strongly linked to an expressive function, and future research may explore to what extent they are related to the emotions of 'sorrow' (*jale*) that these tunes are supposed to convey.

Results also highlighted the higher presence of posterior (or positive) asynchronies over anterior asynchronies (negative), which means that the viola is more often after the violin. Current research has intended the predominance of one type of asynchrony as a proof of clear leader-follower relations. Following Keller (2014), the average asynchrony will be close to 0 ms when there is no prominent lead

part. In our case, the average of asynchrony is always largely greater than 0, which would suggest that the violin is leading. This is certainly true, but more ethnographic data are needed in order to understand how leader-follower relations may have an impact on asynchronies.

Transylvanian musical ensembles are highly hierarchical: the violinist, also called *primas* ('the first') is the band leader. He deals the contract with his clients, chooses the tunes to play at each moment of the performance, sets the tempo, and so on. When they talk about their accompanists, band leaders may explicitly say: 'They are all behind me'. This expression may refer not only to the physical disposition in space, but also to the musical domain: accompanists have to follow and 'wait' the leader, in order to let him lengthening, shortening, and ornamenting the notes in function of his personal style. This leader-follower relation is clear in other musical genres where the degree of improvisation is higher, such as in *doine*, slow tunes in a non-measured rhythm, which are also strongly associated to *de jale* ('of sorrow') emotions. Here, the violinist may introduce new melodic figures in order to assert his personal style, and thus to distinguish himself from other musicians. While playing *doine*, the viola player does not always know in advance 'where the violinist is going', i.e. which melodic contour he is following. So, the viola player needs to hear the notes played by the violinist, before choosing the correct chord to play. This is why posterior asynchronies are generally much more common than anterior asynchronies. We could surmise that this way of playing is also present in the musical genre analyzed here (Gypsy songs of sorrow), even if these tunes are less improvised.[17] Moreover, considering the impressive size of the repertoire which musicians have to know by heart (we are here in a context of oral tradition), playing after the melody is for the accompanist an important mnemonic simplification, since he/she is not obliged to memorize the harmonic development of each tune. In fact, many viola players are not able to reproduce the harmonic pattern of a tune without hearing the melody.

Nevertheless, the viola player may voluntarily anticipate the violinist, especially when he/she knows where the melody 'goes'. These anticipations produce an effect of surprise and of harmonic tension. The liberty of the viola player to anticipate the violinist depends on his social status, his personal relation with the band leader, and his musical quality. Respected viola players may 'impose' their timing pattern to the violinists, shortening and lengthening chords durations in order to keep parts separate. Indeed, while current research on musical ensemble coordination advanced that vertical timing deviations are applied systematically in expressive performances (Shaffer, 1984; Keller, 2014), our results suggest that Transylvanian Gypsy musicians generate both anterior and posterior asynchronies in a rather unpredictable way. Having observed many different musicians playing together, we may affirm that a good performance is the one in which a certain degree of equilibrium is found, that is, when both violinist and violist take the liberty of altering the notes' durations in order to produce this 'out-of-synch' effect that constitute the aesthetic of the Gypsy songs of sorrow. In our opinion, this explains why different type of asynchronies may be present in different performances of the same song, and why the *aksak* S:L proportions may largely vary in the course of a same performance.

FINAL REMARKS

Analysis and transcription of non-western music has always been a controversial subject in ethnomusicology (England, 1964). The Western musical theory and notation system are often unsuited to describing musics of oral traditions which rely on different concepts of rhythm, scale, interval, etc., which are often implicit. For this reason, ethnomusicologists have progressively developed specific techniques in order to analyze and represent the formal properties of the music they study (Rouget, 1981). The main idea underlying the present article is that motion capture techniques can be particularly useful to analyze repertoires of oral tradition, such as Transylvanian village music. Specifically, they were used here to analyze two parameters of the Transylvanian Gypsy 'songs of sorrow' repertoire: *aksak* rhythm and the synchronization between melody and harmonic-rhythmical accompaniment. The tracking of two markers only, positioned at the top of the violin and viola bows, enabled us to analyze the S:L proportion of the *aksak* pattern and the different types and degrees of asynchrony with a higher precision than the measures obtained from the sound signal. While it is widely accepted that 'embodied' representations of the musical activity may offer new possibilities for a deeper understanding of music production and perception (among others, see Leman, 2008; Danielsen, 2010; Burger, 2013; Maidhof *et al.*, 2014), we believe that they may also serve as a support for music analysis and transcription, thus contributing to an emerging non-ethnocentric empirical musicology.

Acknowledgements

Filippo Bonini Baraldi conducted field research in Transylvania thanks to a CNRS (French National Center of Scientific Research) doctoral grant and a 'E. Fleischmann' post-doctoral grant of the Ethnology Society (Nanterre, France). The laboratory sessions were held at INSERM (U1093 Cognition, Action and Sensorimotor Plasticity), Université de Bourgogne (France), thanks to financial support from the CNRS 'Music, Cognition, Societies' scientific network. The drawing up of this article was possible thanks to a CREM-LESC financial support and a CAPES Visiting Professor fellowship (PVE) at Federal University of Paraíba (UFPB, Brazil). The authors are grateful to Yves Ballay and Yannick Gerard for the computation of the motion capture data, to Florent Manneveau for the help with the musical transcriptions, to Philippe Jobet for the realization of the interactive animations, to Laurence Fayet for her revision of the English text, to Speranța Rădulescu and Justin London for their commentaries on previous versions of this manuscript.

NOTES

[1] Filippo Bonini Baraldi is also member of the Centre de Recherche en Ethnomusicologie (CREM-LESC), CNRS UMR 7186, Université Paris Ouest Nanterre la Défense. Correspondence can be addressed to: Filippo Bonini Baraldi, Instituto de Etnomusicologia, Faculdade de Ciências Sociais e Humanas, Universidade Nova, Av. Berna 26 C, 1069-061 Lisboa, Portugal. E-mail: filibb@gmail.com.

[2] Thierry Pozzo is also researcher at the Department of Robotics, Brain and Cognitive Sciences, Istituto Italiano di Tecnologia, Genova, Italy.

[3] Filippo Bonini Baraldi conducted a long-term ethnographical field research on the topic of musical emotion in the village of Ceuaș (central Transylvania), from 2004 to 2008. The socio-cultural context and the musical performances of Gypsy musicians are described and analyzed in detail in Bonini Baraldi (2008, 2009, 2013).

[4] Musical groups in this region are traditionally made up of a violin, a *contră* (a specially prepared viola with three strings fitted over a bridge with a flattened curve, tuned A-D-G) and a three-string double bass played with the bow.

[5] Since *aksak* involves two distinct beat classes (S and L) which may not be related to a common underlying pulse, in this article we will indicate the tempo of the S rhythmic unit (corresponding in our transcription to a quarter note), rather than the tempo of the 8th note.

[6] Arom (2004) put forward that a rhythm is an *aksak* only when it is played at a fast tempo. Yet, in the final part of the same article, the author talks of a Zairean lullaby, characterized by the period 2.2.2.3, as an *aksak*. Even if the author does not indicate the tempo, we can imagine that this lullaby was sung at a slow tempo (as lullabies are generally).

[7] Following London (2012), non-isochronous meters (also known as complex or additive meters), are given by cycles of a prime number of rapid articulations (e.g., 7, 11) or uneven divisions of non-prime cycles (e.g., 9 divided 2+2+2+3).

[8] While Luck & Toiviainen (2006) showed the salience of acceleration as a cue for timing, acceleration data of the marker placed at the top of the bow where not analyzed in our study.

[9] 'Minimum Amplitude Onset' (MAO) and 'Peak Amplitude Onset' (PAO) measures could not be obtained for Duo14, since in the Dijon session the violin was recorded in the same channel than the viola.

[10] In the case of MC measures, the mean S:L value was calculated on 11 periods, the last one being excluded since the musicians' movements did not allow to obtain a reliable measure. It should be acknowledged that all MC measures were obtained on short musical excerpts, which do not allow to make robust average calculations.

[11] More precisely, in the 8 tunes used in part 2 of the current study, the violist never plays eighth notes or sixteenth notes, while in the tune CD this happens only twice (periods 5 and 9, see Figure 2).

[12] Of course, in the Western (classical) performance *accelerando* or *rallentando* are not the *only* expressive devices that modify the tones' durations, and timing may be affected at a local level, depending on the structural details of the music (e.g., Repp, 1992).

[13] A full analysis of the musical concept of 'sweetness' is presented in Bonini Baraldi (2015).

[14] A qualitative analysis of asynchronies was also conducted on the audio recordings CD and CM, which are different takes of the same tune (as Duo12, Duo13 and Duo14), played by the same performers one week earlier (CM) and played by different performers 10 years earlier (CD). This analysis suggests that different musicians introduce different types of asynchronies while playing the same tune. Again, these differences may be attributed to the ornamentation of the melodic line, which is different from one musician to another and is locally conceived as a mark of personal style (see the two transcriptions of the same melody played by two different violinists: Figure 2 and Figure 7). Asynchronies in CM and CD were detected only qualitatively, and further research is needed to precisely measure the impact of personal style on musicians' synchronization.

[15] Of course, this does not mean that asynchrony plays a more important role in oral tradition repertoires than in written repertoires, since there are many examples of oral tradition musics in which performers seek perfect temporal synchrony, and vice versa, of written tradition musics in which performers seek temporal asynchrony.

[16] Gypsy musicians use the term *balanzo*, which we could translate with 'balance', when they want to express the optimal rhythmical quality of a tune. This term can be compared to the idea of 'swing' or 'groove' in jazz. Nevertheless, the term *balanzo* is also used in relation to dance tunes, where melody and accompaniment are much more synchronized, and could therefore point to a combination of other musical parameters, such as tempo, attack and articulation.

[17] More precisely, the violinist may 'improvise' at another level, that of ornaments, passing notes, etc.

REFERENCES

Arom, S. (1992). À la recherche du 'temps' perdu: Métrique et rythme en musique. In J.- J. Wunen-Burger (Ed.), *Les Rythmes* (pp. 195-205). Paris: L'Harmattan.

Arom, S. (2004). L'aksak: Principes et typologie. *Cahiers de Musiques Traditionnelles*, 17, 11–48.

Beaudet, J-M., Martínez, R., Rappoport, D., & Salivas, P. Composing multiplicity. Unpublished paper, Centre de recherche en ethnomusicologie (CREM-LESC), Université Paris Ouest Nanterre la Défense, France.

Benadon, F. (2006). Slicing the beat: Jazz eighth-notes as expressive microrhythm. *Ethnomusicology* 50(1), 73–98.

Bonini Baraldi, F. (2015). La douceur, critère d'appréciation musicale chez les Tsiganes de Transylvanie. *Cahiers d'Ethnomusicologie* 28, 23-41.

Bonini Baraldi, F. (2013). *Tsiganes, musique et empathie*. Paris: Éditions de la Maison des Sciences de l'Homme.

Bonini Baraldi, F. (2009). All the pain and joy of the world in a single melody. A Transylvanian case-study on music and emotion. *Music Perception*, 26(3), 257-261.

Bonini Baraldi, F. (2008). The Gypsies of Ceuaș, Romania: An 'emotional minority'? In R. Statelova, A. Rodel, L. Peycheva, I. Vlaeva & V. Dimov (Eds.), *The human world and musical diversity: Proceedings from the fourth meeting of the ICTM study group 'music and minorities' in Varna, Bulgaria 2006* (pp. 255-261). Sofia: Bulgarian Academy of Science.

Bonini Baraldi, F., Rodà, A. & De Poli, G., (2006). Communicating expressive intentions with a single piano note. *Journal of New Music Research*, 35(3) : 197-210.

Bouët, J. (1997). Pulsations retrouvées: Les outils de la réalisation rythmique avant l'ère du métronome. *Cahiers de Musiques Traditionnelles*, 10, 107-125.

Bărbăiloiu, C. (1973). *Problèmes d'ethnomusicologie*. (A. L. Lloyd Trans.). *Problems of ethnomusicology*. Cambridge: Cambridge University Press, 1984.

Burger, B. (2013). *Move the way you feel: Effects of musical features, perceived emotions, and personality on music-induced movement*. Unpublished doctoral dissertation, University of Jyväskylä, Finland.

Butterfield, M. (2010). Participatory discrepancies and the perception of beats in jazz. *Music Perception*, 27, 157–176.

Clayton, M. (2000). *Time in Indian music. Rhythm, metre, and form in North Indian rag performance*. Oxford: Oxford University Press.

Cler, J. (1994). Pour une théorie du rythme aksak. *Revue de Musicologie*, 82, 181-210.

Cler, J. (1998). Le sens du rythme, une approche ethnomusicologique. In C. Méchin, I. Bianquis & D. Le Breton (Eds.), *Anthropologie du sensoriel. Les sens dans tous les sens* (pp. 107-121). Paris: L'Harmattan.

Cler, J. & Estival, J. P. (1997). Structure, mouvement, raison graphique: Le modèle affecté. *Cahiers de Musiques Traditionnelles*, 10, 37-80.

Danielsen, A. (Ed.) (2010). *Musical rhythm in the age of digital reproduction*. Farnham, Surrey: Ashgate.

England N. M. (1964). Symposium on transcription and analysis: An Hukwe song with musical bow. *Ethnomusicology*, VIII (3), 223-277.

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

Gabrielsson, A. (1995). Expressive deviations and performance. In R. Steinberg (Ed.), *Music and the Mind Machine* (pp. 35-47). Berlin: Springer-Verlag.

Goebel, W. (2001). Melody lead in piano performance: Expressive device or artifact? *Journal of the Acoustical Society of America*, 110, 563–572.

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

Goldberg, D. (2014). Diversity of performance timing in Balkan folk music. Paper presented at the Third International Conference on Analytical Approaches to World Music (AAWM), July 1-4, SOAS, University of London.

Haplea, D. (2005). *Folclor instrumental românesc din Transilvania. Ritmuri asimetrice și tipare modale*. Cluj-Napoca: Arpeggione.

Hood, M. (1982). *The Ethnomusicologist*. (2nd ed.). Kent: The Kent State University Press.

Hove, M. J., Keller, P. E., & Krumhansl, C. L. (2007). Sensorimotor synchronization with chords containing tone-onset asynchronies: The role of P-centers. *Perception and Psychophysics*, *69*, 699–708.

Jankowsky, R. C. (2013). Rhythmic elasticity and metric transformation in Tunisian Štambēlī. *Analytical Approaches to World Music*, *3*(1). retrieved from http://www.aawmjournal.com/articles/2014a/Jankowsky_AAWM_Vol_3_1.pdf

Johansson, M. (2009). *Rhythm into style. Studying asymmetrical grooves in Norwegian folk music*. Unpublished doctoral dissertation. University of Oslo, Norway.

Keil, C. (1995). The theory of participatory discrepancies: A progress report. *Ethnomusicology*, *39*, 1–19.

Keller, P.E. (2014). Ensemble performance: Interpersonal alignment of musical expression. In D. Fabian, R. Timmers, & E. Schubert (Eds.), *Expressiveness in music performance: Empirical approaches across styles and cultures* (pp. 260-282). Oxford: Oxford University Press.

Keller, P. E., & Appel, M. (2010). Individual differences, auditory imagery, and the coordination of body movements and sounds in musical ensembles. *Music Perception*, *28*, 27–46.

Kolinski, M. (1973). A cross-cultural approach to metro-rhythmic patterns. *Ethnomusicology*, *17*(3), 494–506.

Kubik, G. (1988). Einige Grundbegriffe und Konzepte der Afrikanischen Musikforschung. In C. Kaden (Ed.), *Zum Verstehen Afrikanischer Musik* (pp. 52-106). Leipzig: Verlag Philipp Reclam.

Kvifte, T. (2007). Categories and timing: On the perception of meter. *Ethnomusicology*, *51*(1), 64–84.

Leman, M. (2008). *Embodied music cognition and mediation technology*. Cambridge, Massachusetts, USA: MIT Press.

Loehr, J. D., & Palmer, C. (2011). Temporal coordination between performing musicians. *Quarterly Journal of Experimental Psychology*, *64*, 2153–2167.

London, J. (2012). *Hearing in time: Psychological aspects of musical meter*. New York and Oxford: Oxford University Press.

Lucas, G., Clayton, M., & Leante, L. (2011). Inter-group entrainment in Afro-Brazilian Congado ritual. *Empirical Musicology Review*, *6*, 75–102.

Luck, G., & Toiviainen, P. (2006). Ensemble musicians' synchronization with conductors' gestures: an automated feature-extraction analysis. *Music Perception*, *24*, 189-200.

Maidhof, C., Kästner, T. & Makkonen, T. (2014). Combining EEG, MIDI, and motion capture techniques for investigating musical performance. *Behavior Research Methods*, *46* (1), 185-95.

Marchini, M., Papiotis, P., & Maestre, E. (2012). Timing synchronization in string quartet performance: a preliminary study. In *Proceedings of the International Workshop on Computer Music Modeling and Retrieval (CMMR12)*. London, UK.

Martínez, R. (1996). El Sajjra en la música de los Jalq'a. In M. P. Baumann (Ed.), *Cosmología y música en los Andes* (pp. 311-322). Berlin: International Institute for Traditional Music, Vervuert Ibero-Americana.

Moore, G., & Chen, J. (2010). Timings and interactions of skilled musicians. *Biological Cybernetics*, *103*, 401–414.

Patel, A.D., Lofqvist, A. & Naito, W. (1999). The acoustics and kinematics of regularly timed speech: A database and method for the study of the p-center problem. In J. J. Ohala (Ed.), *Proceedings of the XIVth International Congress of Phonetic Sciences* (pp.405-408). San Francisco.

Palmer, C. (1989). Mapping musical thought to musical performance. *Journal of Experimental Psychology: Human Perception and Performance*, 15, 331–346.

Pantaleoni, H. (1972). Three principles of timing in Anlo drumming. *African Music*, 5(2), 50–62.

Polak, R. (2010). Rhythmic Feel as Meter: Non-Isochronous Beat Subdivision in Jembe Music from Mali. *Music Theory Online* 16(4).

Polak, R. & London, J. (2014). Timing and meter in Mande drumming from Mali. *Music Theory Online*, 20(1). Retrieved from <http://www.mtosmt.org/issues/mto.14.20.1/mto.14.20.1.polak-london.php>.

Rappoport, D. (1999). Chanter sans être ensemble. Des musiques juxtaposées pour un public invisible. *L'Homme*, 152, 143-162.

Repp B. H. (1992). Diversity and commonality in music performance: an analysis of timing microstructure in Schumann's 'Träumerei.' *Journal of the Acoustic Society of America* 92, 2546–2568.

Repp, B. H. (1996). Patterns of note onset asynchronies in expressive piano performance. *Journal of the Acoustical Society of America*, 100, 3917–3932.

Rouget, G. (1981). Ethnomusicologie et représentations de la musique. *Le Courrier du CNRS, hors série* 42, 10-11.

Sárosi, B. (1978). *Gypsy music*. Budapest: Corvina Press.

Seeger, A. (1987). *Why Suyà sing. A musical anthropology of an Amazonian people*. Cambridge: Cambridge University Press.

Shaffer, L. H. (1984). Timing in solo and duet piano performances. *Quarterly Journal of Experimental Psychology*, 36A, 577–595.

Stone, R. (1985). In search of time in African music. *Music Theory Spectrum*, 7, 139–48.

Temperley, D. (2000). Meter and grouping in African music: A view from music theory. *Ethnomusicology*, 44(1), 65–96.

Wing A. M., Endo S., Bradbury A., Vorberg D., (2014). Optimal feedback correction in string quartet synchronization. *J. R. Soc. Interface* 11, no. 9320131125.

DISCOGRAPHY

Bolivie. Musiques calendaires des vallées centrales. (1992). CD, Le Chant du Monde (LDX 274938), track 5. Paris: CNRS-Musée de l'Homme.

Szászcsovás Band 3. (1998). CD, Thermal Comfort Kft (3BZATH2), track 6. Budapest, Hungary.