# An experimental validation of Temporal Semiotic Units and Parameterized Time Motifs — Source link [↗]

Aline Frey, Xavier Hautbois, Philippe Bootz, Charles Tijus

**Institutions:** University of Paris, Centre national de la recherche scientifique

Related papers:

- Audio-based Music Structure Analysis

- Linear Basis Models for Prediction and Analysis of Musical Expression

- A Study of Musical Features for Melody Databases

- Multi-scale and multi-dimensional modelling of music structure using polytopic graphs

- Musical similarity in a polyphonic context : a model outside time

# An experimental validation of Temporal Semiotic Units and Parameterized Time Motifs

Aline Frey, Xavier Hautbois, Philippe Bootz, Charles Tijus

# An experimental validation of Temporal Semiotic Units and Parameterized Time Motifs

**Aline Frey**
Laboratoire Cognitions Humaine et Artificielle, Université Paris, France

**Xavier Hautbois**
Centre d'Histoire Culturelle des Sociétés Contemporaines, Université de Versailles-St-Quentin, CNRS, France

**Philippe Bootz**
Laboratoire Paragraphe, Université Paris, France

**Charles Tijus**
Laboratoire Cognitions Humaine et Artificielle, Université Paris, France

## Abstract

Nineteen Temporal Semiotic Units (TSU) have been defined by musicians and researchers from the Laboratoire Musique et Informatique de Marseille (MIM) as categories of musical segments that acquire meaning through their temporal organization. The model of Parameterized Time Motifs (PTMs) provides a graphic modelling of TSUs, which is based on temporal functions and allows for their synthesis. The aim of the present study is to validate the model of PTMs through categorization. To do this, musician and non-musician participants were made to listen and group together three kinds of musical segments: TSUs taken from existing musical works, musical segments played on the piano that correspond to TSUs, and TSUs synthesized from the PTM model. Results have shown that each musical segment, whether synthesized from the PTM model or not, was more frequently paired with other segments of the same TSU category than with other TSU segments. No strong effect of musical expertise was observed. Finally, some TSU categories seemed to be more easily identifiable than others. The significance of these results is discussed in regard to the categorization theory and the effect of implicit learning on perception and knowledge.

**Corresponding author:**
Aline Frey, Laboratoire CHArt, EPHE, 41, rue Gay Lussac, Paris, 75005, France.
Email: aline.frey@cognition-usages.org

## Introduction

The primary task of the auditory system is to arrange the cacophony of frequencies into meaningful clusters that correspond to various real-world activities (Huron, 1991). In the field of music perception, that means that the listener simultaneously perceives the roll of the drums, crash of the cymbal, and brief pulse of noise from the woodwinds as a single coherent event with its own striking properties (Bregman, 1990, p. 460). It is well known that bottom-up and top-down processes are involved during the auditory scene analysis (Alain, Arnott & Picton, 2001).

In this article, we argue that the existence of temporal figures guide the auditory perception. This idea has already been suggested by a few composers and musicologists (Cohen-Lévinas, 2005, pp. 36–37; Grabócz, 2008; Imberty, 1981; Kramer, 2004; Petitot, 1989). In visual perception literature, there are basic geometrical figures (such as squares, circles and triangles) that are seen as primitive units that allow for the decomposition of space (Milner, 1974), either by matching templates (Tarr, 1995), through decomposition into primitive contours, named codons (Hoffmann & Richards, 1984), or by using geon structural descriptions (Biederman, 1987). To our knowledge, there is no such basic sonometric figure equivalent to musical figures in auditory perception literature, let alone in musical flows. As an "art of time" (Francès, 1958), music lacks singular figures that would be equivalent to geometric figures in space. However, Dehaene, Izard, Pica and Spelke (2006) have highlighted how geometrical figures are a concept inherent to the human mind and emphasized their important contribution in the segmentation and understanding of the great diversity of the visual world.

Pierre Schaeffer (1966), a French researcher, theorist, and composer, was the first person to systematically explore this idea and introduce the notion of sound object. He called a sound object every sound event that, heard in a "limited listening," is perceived as a coherent whole, independently of its origin or meaning. A sound object represents a global perception, unchanging from one listening to another, an organized whole that can be assimilated to a *gestalt*. Therefore, Schaeffer has classified sound objects according to their spectrum and temporal evolution through a typology based on seven criteria: mass, dynamics, timbre, melodic profile, mass profile, grain, and look. However, Schaeffer has acknowledged the limits of his experimental project, namely that sound objects are morphological units standing out from a sound flow only under the *gestalt* laws; he has never tackled the issue of their semantic content.

From this point of view, the Temporal Semiotic Units (TSUs), following from Schaeffer's work, represent an attempt to give a consistent definition to a small number of temporal figures for which, contrary to Pierre Schaeffer's sound objects, "we look for the smallest segment corresponding to a defined meaning" (Delalande et al., 1996, p. 19).

## Temporal Semiotic Units (TSUs)

The origins of the TSUs derive from an observation made by musicians and researchers from the Laboratoire de Musique et d'Informatique de Marseille (MIM) in 1991. While Schaeffer aims to describe sound through a morphological description based on a "limited listening" (to quote Schaeffer's term), which consists in ignoring any causal or associative meaning, the composition of an electroacoustic piece appears to be a long way from this practice. For example, French composers such as Pierre Henry or François Bayle select their sounds and classify them according to their meaning, based on an associative denomination, on what the sound "says" (i.e., "something in there had to seem to be singing to me," Bayle, quoted in Delalande,1990, pp. 51–65).

Therefore, the MIM researchers have proposed to re-instill meaning in their description of sound objects and define some types of *semiotic objects*, called Temporal Semiotic Units. It should be emphasized that the MIM's work is consistent with various empirical and theoretical approaches to classification based on time perception, used by composers such as Salvatore Sciarrino (1998, *figore*), Trevor Wishart (1994, *morphology of sounds*), Denis Smalley (1986, *spectro-morphology*), François-Bernard Mâche (2001, *archetypes/genotypes/phenotypes*), Costin Miereanu (1996, *elementary formal categories*) and François Bayle (2003, *morphodynamic categories*). For all these composers, categories stem from their own auditory and compositional experience.

Assuming that some musical figures seem to produce both a specific temporal meaning and kinetic effects, the research group looked for any remarkable dynamic or static sound effect (uniform or accelerated motion, circular or linear process, compression, propulsion or stationary state, etc.) in a number of electroacoustic pieces. Let us consider the effect of the illusions induced in the auditory system by Roger Shepard, as illustrated by Jean-Claude Risset in *Mutations* (2001): it is a sound whose pitch continually increases or decreases and that ultimately does not seem to get higher or lower. These sounds have a precise single-stage temporal evolution. We are not interested in the sound illusion here, but in the inexorable and predictable process related to the pitch parameter. This temporal process corresponds to a unified sound figure characterized by the MIM researchers as *Endless trajectory*. Another example is a two-phase sound figure frequently encountered in Bernard Parmegiani's works (*De natura sonorum*, 2001), which is called *Compressing–stretching out*: the first fairly short phase including small accumulated and compressed sound elements is followed by a second period, opposite to the first and containing a consistent sound that slowly extends over time. Of course, this figure can be described in purely morphological terms but the impression that we get from the accumulation of sounds followed by an expansion over time, two phases constituting a coherent and complete unit, is exclusive to this temporal figure. Some morphological criteria must be observed to produce the effect (they are relevant morphological traits) while others do not make any difference (sound figures can be played in any register: low, middle or high), but it is clearly in the temporal meaning of the *Compressing–stretching out* or *Endless trajectory* that we can identify and isolate a unit. In this respect, they are Temporal Semiotic Units. TSUs have been defined as "musical segments that possess a precise temporal signification linked to their morphological organization" (Delalande et al., 1996, p.18). In other words, TSUs are sound forms that convey meaning through the dynamic evolution of their pattern over time. TSUs are not Schaefferian sound objects because they are not isolated from their context according to the same segmentation criteria (that is why the name "semiotic object" was finally dropped). The Schaefferian sound object is extracted from its context in accordance with pure *gestalt* laws, rather than with meaning units as is the case with TSUs (in which a minimal segment corresponding to a well-defined meaning is sought). Thus the division of sound objects and that of TSUs do not necessarily coincide: a TSU can be made up of a set of sound objects.

In order to observe and list the different TSUs, the MIM researchers first listened to a number of musical works before dividing them into musical segments that have a precise temporal meaning. These segments last about a few seconds and their duration is controlled through working memory (Baddeley, 1986). They are based on a listening attitude that focuses neither on the pitch of a note, nor on the musical harmony or timbre of sounds, but on the overall effect of all these parameters over time. A TSU is a meaning base unit at the end of the segmentation process. Thus, there is no reason to believe that further segmentation (i.e., another listener who performs the segmentation or the same listener who does it several times) would produce other TSUs; similarly, no TSU can be divided into smaller parts that would convey meaning. The

**Table 1.** Identification cards of the TSUs *Endless trajectory* and *Compressing–stretching out*, presenting morphological and semantic descriptions, as well as other relevant characteristics to each defined TSU.

| Endless trajectory | Compressing–stretching out |
|---|---|
| **Morphological description**: One-phase unit, non-time-delimited, with linear and usually slow evolution of a sound parameter.<br>**Semantic description**: The process must be oriented in a direction (for example, upwards or downwards) and yet it seems never to end.<br>**Other relevant required characteristics**: The duration of the sound phenomenon must be as long as to be perceived as a process and not an ephemeral event. | **Morphological description**: Time-delimited unit with two contrasting phases. During the "compressing" phase, the sound matter is discontinuous and erratic. The "stretching out" phase is a globally uniform segment.<br>**Semantic description**: First, there is a feeling of compression (as if we pressed down hard on an obstacle), then the barrier is suddenly overcome, suppressing all resistance and releasing the power. It is a sudden shift from localized energy to scattered energy.<br>**Other relevant required characteristics**: The sound matter of the first phase must have a rougher mass and a thicker grain than the second. Its dynamic profile is a quick crescendo, a maximum intensity or a mix of various sforzandi. The second phase must be deliberately sustained, not a mere resonance. The trajectory between the two phases is not continuous: there is a break. The relative durations of the two phases must allow for the perception of contrast. |

researchers have gathered segments with a similar temporal organization: each category has been described on morphological and semantic levels identifying relevant common characteristics. Four morphological characteristics have been used: duration, which can be time-delimited (these are usually multi-phase TSUs with a distinct beginning and end) or not (in which case, a portion of the unit is enough to identify the whole segment, like the *Endless trajectory*); reiteration (with or without cycles); number of phases (one or more); and sound matter (continuous or discontinuous sound). Two kinetic characteristics have been added to these: the type of acceleration (positive or negative) and temporal progression (fast or slow). Several semantic characteristics are also provided, such as the process direction (one or more sound parameters moving in the same direction); movement (motion effects); and sound energy (constant or retained). Every TSU has been analyzed according to these characteristics and given identification cards as presented in Table 1.

The MIM researchers have defined 19 TSUs so far. They have metaphorically named them as follows: *Braking, Chaotic, Compressing-stretching out, Divergent, Endless trajectory, Fading away, Falling, Floating, Heaviness, In suspension, Moving forward, Obsessive, Propulsion, Spinning, Stationary, Stretching, Suspending-questioning, Wanting to start, Waves*. The Appendix provides a list of the TSUs along with their general description.

It should be clarified that the research described here is in progress, and that the solutions proposed herein should be considered as provisional. Indeed, since TSUs were empirically obtained from contemporary pieces of music (in particular electroacoustic music), there is no guarantee that no other TSU will be discovered in other works yet to be analyzed, works taken from different musical styles or other musical cultures. A few TSUs are especially adapted to music from the second half of the 20th century: such is the case, for instance, of the TSU

**Figure 1.** Score extract from Beethoven's "Geister" Trio first movement (*Allegro vivace e con brio*), representing the TSU *Compressing–stretching out*.

*Divergent* (characterized by a succession of brief temporal sequences bearing no relationship with one another) or the TSU *Chaotic* (which is a dense superimposition of musical sequences resulting in high saturation). However, the MIM researchers have rapidly discovered that most TSUs had a more global scope for the analysis of varied repertoires. Again, let's go back to the two TSUs previously described in order to give examples likely be found in a musical score. In Beethoven's "Geister" Trio, the introductory sequence of the first movement (*Allegro vivace e con brio*) can be analyzed in temporal terms with the TSU *Compressing–stretching out* (Figure 1). The same TSU occurs in the following measures (Figure 2) of the contemporary quartet *Thus the night* by Henri Dutilleux, although they are expressed in a very different musical style.

In Shepard-Risset's work, an *Endless trajectory* can be observed in this slowly decreasing melodic line (Figure 3) from Claude Debussy's prelude "The Terrace of Moonlit Audiences." It is also to be found in the enlarged process of Gyorgy Ligeti's *Devil's Staircase*, which seems to keep rising (Figure 4).

It must be emphasized that, if a music score includes a symbolic time codification, the TSUs are to be perceived, or not, in the musical performance of the interpreters. If the pianist plays Debussy's prelude a little bit too fast, marking the four-note motif of the left-hand octaves (instead of following the composer's clear indications: "a bit on the outside"), the feeling conveyed by this *Endless trajectory* is likely to escape us. Likewise, if the three musicians of Beethoven's *Allegro vivace con brio* play the first bars of the score without enough *brio*, that is if they are not energetic enough, the perception of the *Compressing–stretching out* TSU may not be relevant. TSUs allow for an analysis without a score, without any prerequisite knowledge, because it is an analysis "beyond notes" based on sound perception (Delalande, 2008, p. 17). Surely, temporal semiotics is not the only semiotics to be applied to music and the TSU analysis sheds so particular a light that it can only partially cover a musical work (like any other method of analysis).

Therefore, originally created for electroacoustic music, for which traditional vocabulary is sometimes inappropriate, the TSUs have proven to be efficient in France as an analytical tool (in

**Figure 2.** Score extract from the Henri Dutilleux's contemporary quartet *Thus the night*, representing the TSU *Compressing–stretching out*.
Reprinted with kind permission and copyright of Heugel S.A. Paris



**Figure 3.** Score extract from the Claude Debussy's prelude "The Terrace of Moonlit Audiences,"
representing the TSU *Endless trajectory*.
Reprinted with kind permission of MGB Hal Leonard.

works by Guy Reibel, Henri Dutilleux, Claude Debussy, Arvo Pärt, Luciano Berio, Costin Miereanu, Ivo Malec, Franco Donatoni, Jean-Pierre Moreau, Francis Dhomont, etc.; cf. Favory et al., 2002), as an educational tool (TSUs are studied in a class at Marseille's Regional Conservatory, in schools in the Academies of Montpellier, Grenoble and Besançon, the Municipal Music School of Cabriès, etc.) and has also become a compositional tool (Di Santo, 2008). The TSUs are now widely recognized, as expressed by Landy (2007) who emphasized the relevance of the project, they provide an analytical tool that is efficient both on a morphological and semantic level. We share the view of Austin (2011) who suggests that the use of TSUs as an analytical technique should be extended to sonic art and multimedia. Moreover,

**Figure 4.** Score extract from Gyorgy Ligeti's *Devil's Staircase*, representing the TSU *Endless trajectory*. Reprinted with kind permission of Schott Music, Mainz, Germany.

Pierre Couprie's musical analysis software EAnalysis, which is part of the research project entitled "New multimedia tools for electroacoustic music analysis" (funded by the Arts and Humanities Research Council (AHRC)) at De Montfort University MTI Research Centre, Leicester, UK), resorts to TSUs (Couprie, 2013).

However, in the field of cognitive psychology, TSUs raise some issues. Delalande et al. (1996) emphasized that:

> insofar as the constitution of equivalent examples on semantic criteria rely on the judgment of a group of people, we conceive that this equivalent judgment could be, a priori, relative on one hand to this group, and on the other hand to a kind of attitude or commonly adopted listening behavior. We hope to show that our results could be extended to a wider group and maybe introduce generalities on listeners. They will most likely remain relative to a certain listening attitude and one will have to focus on a specific direction to become aware of temporal semantics. (p. 21)

Indeed, what is the psychological validity of the TSU? Does the perception of the TSU require a listening that would necessarily focus on temporal progression? Likewise, the work of listening, segmenting, defining the TSUs has been carried out by a team of musicians: not only do their competencies benefit our work, they are also important to ascertain to what extent musical expertise is necessary to understand TSUs.

In Frey, Daquet et al. (2009), a set of experiments was conducted to answer these points. In these experiments, participants—both musicians and non-musicians—were asked to group together musical samples derived from different TSU categories. The main results can be summarized as follows: samples derived from the same TSU are grouped together significantly more often than samples derived from other TSUs; moreover, no significant effect of the participants' musical expertise was observed in the number of groups that were formed, or in the

composition of those groups. Eventually, the observed groups are not "better," i.e., corresponding well to TSU categories, through a categorization task in which the experimenter has asked the participants to focus on the temporal progression of samples, in comparison with a free sorting task, performed by musician or non-musician participants alike. The results of those experiments show that TSUs are relevant sound units, for all kinds of listeners, regardless of any particular listening attitude. This constituted a first test on the validity of TSUs. In Frey, Marie, et al. (2009), the temporal meaning of TSUs was assessed through an electrophysiological approach relying on the event-related potentials method to determine whether conceptual priming could occur between two successively presented TSUs. The most interesting result shows that an incongruous TSU, i.e., a TSU that starts as one TSU but shifts midstream to another TSU, elicits an N400-like component in participants who are not experts in TSUs. This component is well known to be a good index of the lack of integration of a stimulus within its preceding semantic context (Kutas & Federmeier, 2000) and reflects cognitive processes associated with semantic expectancy (Kutas & Hillyard, 1984).

Another kind of validation should rely on the operational nature of TSUs. Indeed, every TSU category has its own morphological and kinetic properties and it should be possible to model TSUs, namely to define parameters that evolve according to some forms. This modelling task has been carried out by Bootz and Hautbois (2007, 2008) and led to the Parameterized Time Motifs PTMs.

## The Parameterized Time Motifs (PTM)

The PTMs are an attempt by Philippe Bootz (researcher in media studies) and Xavier Hautbois (musicologist) to analyze TSUs according to the quantitative evolution of sound parameters rather than using qualitative or metaphorical descriptions. They have used analytic functions to describe the temporal progress of TSUs (Bootz & Hautbois, 2007, Bootz & Hautbois, 2010a). A PTM represents the signifier of a TSU, through temporal functions applied to relevant variables, configured on the whole duration of the TSU. The signifier refers to the acoustic image of the stimulus (the sound itself), i.e., an abstract model derived from the stimulus (Klinkenberg, 1996). More specifically, those temporal functions are not about the detailed development of physical dimensions, but rather about the representation of the global evolution of variables. Those variables describe the global evolution of the most perceived changes in temporality.

For example, the third movement of Shostakovich's Symphony No. 8 (*Allegro non troppo*) begins with an archetypal sequence that corresponds to the *Obsessive* TSU. The sound wave and the sonogram show the temporal evolution of the sound intensity and frequencies (Figure 5).

From these temporal representations of sound, we have defined a temporal motif. In our example, the sonogram shows a periodic structure, in which the periodic motif is not strictly identical from one period to another. As a model, a PTM needs not have all the sound characteristics of the corresponding stimulus of the TSU, it will only describe the common sound properties corresponding to an *Obsessive* TSU. The exact value of the motif's period is not essential to model and recognize an *Obsessive* TSU. In the same way, the exact shape of the elementary motif is not relevant and a pattern with an increasing and decreasing curve is sufficient (within such a small period, a symmetric curve is as effective as an asymmetric one and a percussive envelope is also appropriate). For this *Obsessive* PTM, the sound volume is relevant and represented by a bell-shaped time curve on the relevant variable I. Frequency variable is not relevant in the signifier of this TSU and is thus represented by a flat shape. The three axes (F, I, time) of
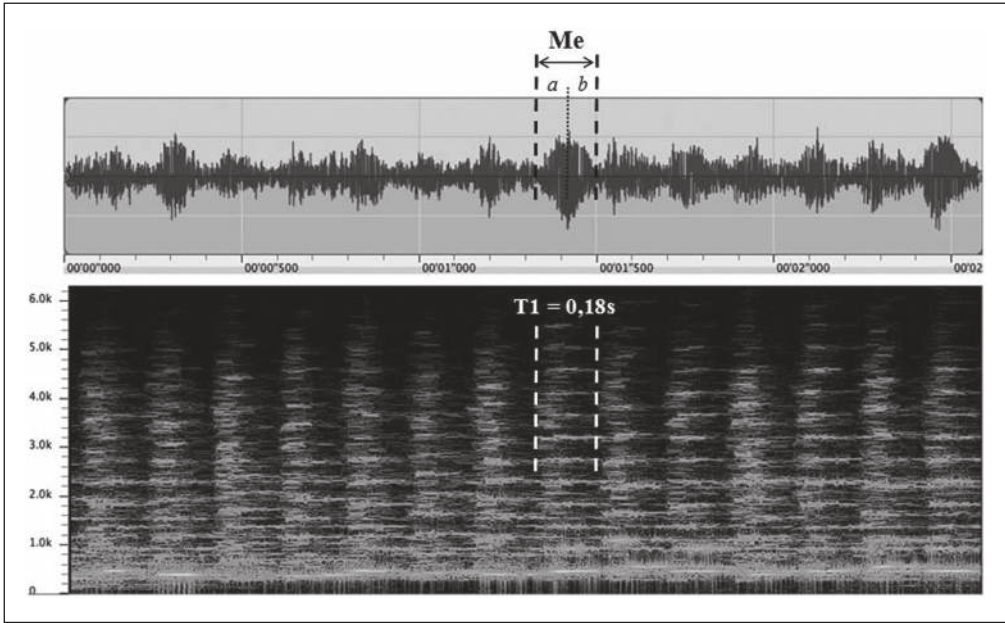
**Figure 5.** Sound waves and sonogram of an extract from Shostakovich's Symphony No. 8.
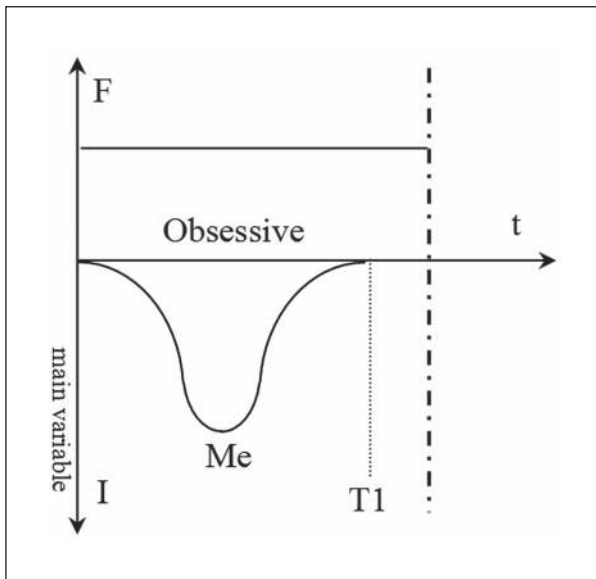


**Figure 6.** PTM of the *Obsessive* TSU. The figure shows an elementary motif Me (a bell-shaped time curve) on the relevant variable I. The curve is repeated with a short period T1 (< 1s). Frequency variable F is not relevant (a flat shape).

this PTM, along with all these relevant properties describing the signifier taken from the musical flow, can be represented in a single conventional figure (Figure 6), in which each and every
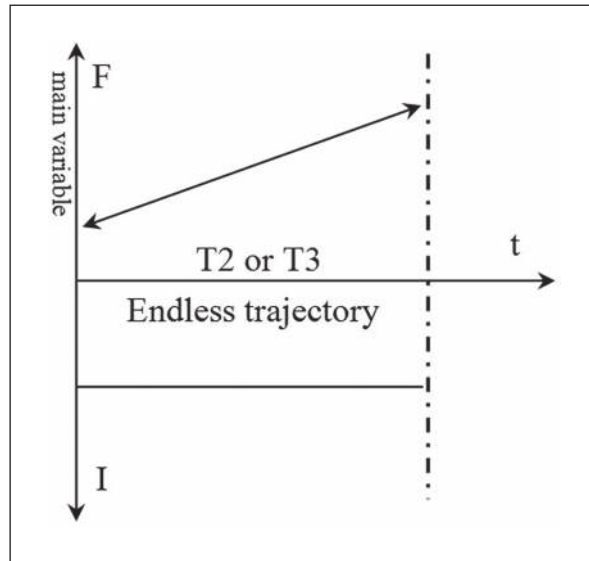
**Figure 7.** PTM of the *Endless trajectory* TSU. In this figure, the relevant variable F follows an increasing or decreasing straight line (represented by the double arrow). The temporal function lasts several seconds (T2 or T3). The variable I is not relevant (a flat shape).

element is a vocabulary component. The indication "Me" points out the elementary motif, which will be repeated in the sound (period T1 < 1s). The end of the PTM is represented by a bold intermittent line. A few terms of vocabulary are needed to describe the very high diversity of physical stimuli of a TSU, and the vocabulary used is the same for all TSUs.

The main part of this vocabulary consists of temporal functions represented in the form of a PTM. A temporal function is constructed when a dynamic effect is perceived in the global sound flow, for example in a whole orchestral behavior (accelerated motion, circular or linear process, etc). Thus, a temporal function could be continuous, even when the real temporal evolution is not. In the aforementioned musical examples of *Endless trajectory*, the relevant variable (pitch linear evolution) is represented by a simple straight line with a double arrow (which means it can go upwards or downwards). The change in intensity is not relevant. The duration of the TSU must be long enough to allow for the perception of the process (T2 or T3 implies a few seconds) (see Figure 7). The temporal function does not reproduce brief and local changes, but is close to an average calculated on a large temporal window and tallied with a kind of macro-structure. It expresses the evolution of a perceived physical parameter (Bootz & Hautbois, 2010a). As a rule, those temporal functions have been represented in a graph, in the form of a time-dependent continuous line.

Two pertinent variables are enough to describe the signifier of a TSU. These abstract variables have been named F and I: the variable F materializes the "melodic" aspect of sound (frequencies evolving over time) and the variable I principally materializes the sound intensity, or in other cases its brilliance, which is to say the spectral richness of sound. Each TSU has been systematically described with these two variables F and I. Nevertheless, the pattern of only one of these variables suffices to perceive the TSU in many cases. Consequently, it has been named "main variable of the TSU." When the TSU does not possess such a variable, the two pertinent variables are of equal importance: the perception is as influenced by the time-pattern of one as it is by the other.

The PTMs can be seen as abstract TSUs, or their temporal skeletons. They allow for the comparison of TSUs, although they could not be easily obtained from the MIM's identification cards. For example, *Waves* and *Obsessive* have similar temporal profiles (bell-shaped time curve on the relevant variable I): the difference between them is the period of the elementary motif (the period of the *Obsessive* TSU is short, contrary to that of the *Waves* TSU). In this sense, a transition from the *Waves* TSU to the *Obsessive* TSU can occur through a variation in the length of the motif's period. The complete comparison between TSUs and the PTM model is in progress.

The PTM model can be used to synthesize sound sequences in a visual programming language such as Open Music (developed by IRCAM, Paris). It would be too long to develop the model in this article: all PTMs are presented in Bootz and Hautbois (2007, 2008), namely the details of the construction of PTMs and their graphic representation. The authors have worked directly from an initial sound file containing a steady and homogeneous sound. They have applied filters to this sound in order to vary frequency and intensity according to the temporal functions of the model. For example, to process an increasing variation of frequencies, starting from an initial sound file containing a homogeneous sound (yet likely to be a complex sound with internal variations) they used a pitch shifter that raises the global pitch of sound. The simulation program has been designed to allow for the easy creation of a vast corpus of sound samples, by playing on the different parameters and choice of the initial sound file.

## The present study

One may still wonder whether PTMs form an appropriate model of TSUs and whether they have a psychological validity as mental categories of sound forms. In particular, the aim of this experiment is to check if the sound segments synthesized from the model of PTMs are perceived as TSUs. To do this, we carried out a categorization task, during which we asked participants to listen to musical segments and group them together.

Categorizing a stimulus amounts to considering it as equivalent to other stimuli and yet different from those which do not belong to its category (Tijus & Cordier, 2003). Within the same category, there is a single relationship between objects, which is different from that between objects in other categories (Dubois, 1991; Rosch & Mervis, 1975). Each categorization task or recognition of a given object can partially be described in terms of an estimation of the similarity between this object and the conceptual representations of the subject (Thibaut, 1997).

In our experiment, musical segments that were presented to participants included segments synthesized from the model of PTMs, musical extracts taken from existing musical works and corresponding to TSUs, as well as short compositions played on the piano and strictly respecting the definition of TSUs. Consequently, the only common characteristic between these three kinds of musical segments was the temporal progression. Thus, we hypothesized that participants would group together the musical extracts according to the TSU they were derived from, independently of the nature of the musical extract (segments synthesized from the model of PTMs, TSUs taken from musical works or TSUs played on the piano). More specifically, we hypothesized that if PTMs provided a good modelling of TSUs, participants should group them together with the other extracts derived from the same TSU category.

The experiment was conducted with both musician and non-musician participants. Numerous studies have shown functional and anatomical differences between those two groups of expertise (Berz, 1995; Deutsch, 1970, 1972a, 1972b, 1973; Pechman & Mohr, 1992; Peretz & Zatorre, 2005; Tervaniemi, Just, Koelsch, Widman, & Schröger, 2005; Williamon & Egner, 2004; Williamon & Valentine, 2002), and we hypothesized that our musician participants would form groups that would match TSU categories better. However, as seen in the

introduction to this article, TSUs allow for an analysis that is independent from any prerequisite knowledge, and we could also hypothesize, as observed in Frey, Daquet et al. (2009), that non-musicians would perform as well as musicians, probably because the "expertise" necessary to perceive TSUs is an expertise linked to the act of listening, common to musicians and non-musicians.

## Method

### Participants

Thirty-two participants, with no hearing impairment, participated in this study: 16 were musicians (mean age = 18.20 years; SD = 2.32) and 16 were non-musicians (mean age = 21.93 years; SD = 5.13). Musicians came from the Ecole Nationale de Musique, de Danse et d'Art Dramatique (ENDMA) in Bobigny, France and the Conservatoire Municipal de Musique, Danse et Art Dramatique Henri Dutilleux in Maisons-Alfort, France. They had a minimum of six years of musical and theoretical tuition. Non-musicians mostly came from the University Paris 8, France. None of the participants from either group knew about the TSUs or the PTMs beforehand.

### Materials

Fifteen TSU categories were used in our experiment, namely *Braking, Compressing–stretching out, Endless trajectory, Fading away, Falling, Floating, In suspension, Moving forward, Obsessive, Propulsion, Spinning, Stationary, Suspending–questioning, Wanting to start, Waves*. The four remaining TSU categories (*Chaotic, Divergent, Heaviness, Stretching*) were not considered because musical extracts were too long or too difficult to play on the piano (because of the difficulty in playing a continuous sound).

Each and every of the 15 TSU categories was represented with three sound segments:

- a musical segment taken by the MIM researchers from the existing musical repertoire (mean duration = 20.62 s; SD = 13.05);
- a segment played on the piano by a composer from the MIM (L. Prod'Homme), strictly complying with the definition of every TSU (Delalande et al., 1996; mean duration = 16.10 s; SD = 10.70);
- a segment resulting from a sound synthesis of the corresponding PTM (mean duration = 14.82 s; SD = 7.61).

Forty-five musical segments were used during the experiment. The sound synthesis of the PTMs was carried out using OpenMusic V 5.0 software from the IRCAM (Institut de Recherche et de Coordination Acoustique / Musique, Paris), designed for computer-assisted composition. The program has been developed by X. Hautbois (Bootz & Hautbois, 2008). Musical examples of TSUs and PTMs are available from: http://www.labo-mim.org/site/index. php?2013/03/29/225-temporal-semiotic-units-tsus-a-very-short-introduction. The software used in the experiment was FRIDA Sons,[1] which allows sound files to be associated with icons, here represented by balls, and move them to a specific area on the screen. Among the 15 TSU categories represented, FRIDA Sons randomly drew out six TSU categories at the beginning of each free-sorting task, to limit the duration of the experiment and the mental effort required from the participant. Then, each participant was confronted with 18 sound

segments, three for each of the six categories of TSU that were randomly displayed on the screen in the form of 18 balls. By clicking on a ball, the participant would listen to the associated sound segment. We used Sennheiser HD 25 SP headphones plugged to an iBook G4.

## Procedure

Each participant individually saw 18 balls on the screen, which corresponded to 18 sound segments. Their task was to group together sound segments that sounded alike. There was no set criterion regarding the constitution of the groups (free categorization), but in the instructions, we encouraged the participants to have an acousmatic listening,[2] namely to focus on the sound effect produced by the segment. The experimenter first recommended that the participants listen to all the sound segments to get acquainted with them and get an idea of their similarities or differences. There was no time pressure and participants were allowed to make as many groups as they wanted. At the end of the experimental session, participants were asked to explicitly show their groups to the experimenter and try to explain what each of them corresponded to.

## Results

Each participant had to group sound segments according to a TSU. As there was a set of three sound segments per TSU and six TSUs, the "ideal" number of groups was six, with three sound segments per group. Participants formed an average of 5.63 groups (SD = 1.39): 5.13 groups (SD = 1.17) for non-musicians, and 6.13 groups (SD = 1.41) for musicians. This difference was significant ($F(1, 30) = 4.49$, $p < .05$): the musicians formed a higher number of groups, which might be related to a higher degree of differentiation. The average number of sound segments per group was 3.41 (SD = 0.87), with 3.71 (SD = 0.89) sound segments for non-musicians and 3.10 (SD = 0.73) for musicians. There was a significant effect of the expertise on the number of sound segments per groups ($F(1, 30) = 4.17$, $p = .05$).

Table 2 represents the average association frequency (for all participants, musicians and non-musicians) of each sound segment in the rows (PTMs, played on the piano and musical segments) with each TSU category in the columns. Association frequency tallied with the number of times that a segment had been associated with the two other segments of its TSU category. For the two TSUs on which we have previously focused, *Endless trajectory* and *Compressing–stretching out*, the results are as follows: the PTM synthesized from the TSU *Endless trajectory*, labeled endless trajectory1, was grouped together 0.35 times with the segments endless trajectory2, played on the piano, and with endless trajectory3 taken from existing musical works. As for the PTM compressing–stretching out1, it was grouped together 0.40 times with the segment compressing–stretching out2 (played on the piano) and with the segment compressing–stretching out3 (musical works).

Note that the probability of a random grouping between a particular segment and one of the two other segments of the same TSU was 2/17 (0.12). We can see on this table that in most cases, a segment was more frequently grouped together with segments of its TSU category than at random, regardless of its type (in bold in Table 2). The average association frequency was 0.30 for PTMs, 0.27 for segments played on the piano and 0.28 for musical segments. The TSUs that were most frequently recognized were: *Compressing–stretching out*, *Propulsion* and *Floating*, with an average association frequency of 0.44, 0.39 and 0.44 respectively. There were only four segments, namely waves1, fading away1, stationary2 and fading away2 for which the association frequency was more important than, or equal to, another TSU compared with the

**Table 2.** Average association frequency for all participants—musicians and non-musicians—of each sound segment (rows) (PTMs played on the piano and musical segments), with each TSU category (columns). Bold cells underline the association frequency of a sound segment with its TSU category. Grey cells underline the association frequencies that are more important than, or equal to, another TSU compared with the TSU category the sound segment comes from.

| | Braking | Compressing–stretching | Endless out trajectory | Fading away | Falling | Floating | In suspension | Moving forward | Obsessive | Propulsion | Spinning | Stationary | Suspending questioning | Wanting to start | Waves |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **PTH** | | | | | | | | | | | | | | | |
| barking1 | **0.26** | 0.02 | 0.00 | 0.07 | 0.05 | 0.12 | 0.02 | 0.09 | 0.09 | 0.12 | 0.02 | 0.00 | 0.09 | 0.02 | 0.00 |
| compr.-stret.1 | 0.15 | **0.40** | 0.00 | 0.10 | 0.15 | 0.05 | 0.00 | 0.10 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.05 |
| endl.-traj.1 | 0.00 | 0.00 | **0.35** | 0.09 | 0.06 | 0.00 | 0.00 | 0.06 | 0.00 | 0.06 | 0.06 | 0.09 | 0.09 | 0.06 | 0.09 |
| fading away1 | 0.06 | 0.06 | 0.13 | **0.22** | 0.03 | 0.00 | 0.06 | 0.09 | 0.00 | 0.03 | 0.03 | 0.22 | 0.00 | 0.03 | 0.03 |
| falling1 | 0.00 | 0.06 | 0.00 | 0.00 | **0.31** | 0.00 | 0.06 | 0.19 | 0.09 | 0.00 | 0.00 | 0.06 | 0.06 | 0.09 | 0.06 |
| Floating1 | 0.13 | 0.03 | 0.03 | 0.03 | 0.00 | **0.40** | 0.00 | 0.10 | 0.00 | 0.13 | 0.03 | 0.00 | 0.03 | 0.07 | 0.00 |
| in suspense1 | 0.00 | 0.03 | 0.00 | 0.14 | 0.03 | 0.00 | **0.28** | 0.07 | 0.00 | 0.00 | 0.14 | 0.07 | 0.14 | 0.07 | 0.03 |
| moving for.1 | 0.03 | 0.03 | 0.05 | 0.08 | 0.07 | 0.07 | 0.03 | **0.21** | 0.10 | 0.07 | 0.05 | 0.05 | 0.02 | 0.10 | 0.05 |
| obsessive1 | 0.12 | 000 | 0.00 | 0.00 | 0.10 | 0.07 | 0.07 | 0.15 | **0.24** | 0.05 | 0.00 | 0.07 | 0.03 | 0.05 | 0.07 |
| propulsion1 | 0.21 | 0.00 | 0.03 | 0.00 | 0.00 | 0.13 | 0.00 | 0.05 | 0.00 | **0.42** | 0.05 | 0.00 | 0.03 | 0.08 | 0.00 |
| spinning1 | 0.03 | 0.00 | 0.11 | 0.03 | 0.03 | 0.08 | 0.08 | 0.16 | 0.00 | 0.11 | **0.26** | 0.03 | 0.00 | 0.08 | 0.03 |
| Stationary1 | 0.00 | 0.00 | 0.05 | 0.11 | 0.05 | 0.00 | 0.05 | 0.03 | 0.03 | 0.00 | 0.03 | **0.47** | 0.03 | 0.03 | 0.13 |
| sus.-guest.1 | 0.12 | 0.04 | 0.12 | 0.00 | 0.08 | 0.04 | 0.04 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | **0.40** | 0.16 | 0.00 |
| wanting to s.1 | 0.00 | 0.00 | 0.02 | 0.00 | 0.02 | 0.08 | 0.04 | 0.13 | 0.04 | 0.08 | 0.06 | 0.12 | 0.04 | **0.27** | 0.12 |
| waves1 | 0.00 | 0.04 | 0.04 | 0.04 | 0.02 | 0.00 | 0.04 | 0.03 | 0.04 | 0.00 | 0.02 | 0.27 | 0.02 | 0.14 | **0.24** |
| **Piano** | | | | | | | | | | | | | | | |
| barking2 | **0.28** | 0.03 | 0.00 | 0.10 | 0.03 | 0.08 | 0.03 | 0.08 | 0.03 | 0.25 | 0.03 | 0.03 | 0.03 | 0.00 | 0.05 |
| compr.-stret.2 | 0.06 | **0.59** | 0.00 | 0.00 | 0.13 | 0.00 | 0.00 | 0.12 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.06 |
| endl.-traj.2 | 0.00 | 0.00 | **0.30** | 0.21 | 0.09 | 0.00 | 0.00 | 0.09 | 0.06 | 0.03 | 0.03 | 0.00 | 0.12 | 0.06 | 0.00 |
| fading away2 | 0.06 | 0.00 | 0.17 | **0.11** | 0.00 | 0.03 | 0.14 | 0.22 | 0.03 | 0.03 | 0.03 | 0.06 | 0.06 | 0.06 | 0.03 |
| falling2 | 0.03 | 0.05 | 0.08 | 0.03 | **0.28** | 0.00 | 0.03 | 0.13 | 0.08 | 0.00 | 0.00 | 0.13 | 0.10 | 0.05 | 0.03 |
| floating2 | 0.12 | 0.00 | 0.00 | 0.00 | 0.00 | **0.50** | 0.03 | 0.12 | 0.09 | 0.09 | 0.03 | 0.00 | 0.00 | 0.03 | 0.00 |
| in suspense2 | 0.00 | 0.00 | 0.00 | 0.10 | 0.03 | 0.00 | **0.26** | 0.06 | 0.00 | 0.03 | 0.13 | 0.13 | 0.03 | 0.10 | 0.13 |
| moving for.2 | 0.10 | 0.04 | 0.04 | 0.08 | 0.10 | 0.04 | 0.03 | **0.15** | 0.11 | 0.01 | 0.08 | 0.04 | 0.01 | 0.08 | 0.08 |

*(Continued)*

**Table 2.** (Continued)

| | Braking | Compressing–stretching | Endless out trajectory | Fading away | Falling | Floating | In suspension | Moving forward | Obsessive | Propulsion | Spinning | Stationary | Suspending questioning | Wanting to start | Waves |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| obsessive2 | 0.05 | 0.00 | 0.02 | 0.02 | 0.05 | 0.07 | 0.05 | 0.20 | **0.27** | 0.05 | 0.00 | 0.10 | 0.00 | 0.05 | 0.07 |
| propulsion2 | 0.13 | 0.00 | 0.06 | 0.03 | 0.00 | 0.03 | 0.03 | 0.03 | 0.09 | **0.38** | 0.03 | 0.03 | 0.03 | 0.13 | 0.00 |
| spinning2 | 0.09 | 0.00 | 0.00 | 0.00 | 0.06 | 0.00 | 0.14 | 0.09 | 0.00 | 0.03 | **0.26** | 0.03 | 0.00 | 0.17 | 0.14 |
| static nary2 | 0.02 | 0.00 | 0.03 | 0.06 | 0.05 | 0.02 | 0.03 | 0.08 | 0.11 | 0.05 | 0.06 | **0.20** | 0.02 | 0.08 | 0.22 |
| sus.-quest.2 | 0.00 | 0.07 | 0.21 | 0.07 | 0.03 | 0.00 | 0.17 | 0.00 | 0.00 | 0.00 | 0.00 | 0.07 | **0.24** | 0.10 | 0.03 |
| wanting to s.2 | 0.03 | 0.02 | 0.07 | 0.05 | 0.10 | 0.00 | 0.08 | 0.11 | 0.00 | 0.05 | 0.08 | 0.02 | 0.08 | **0.18** | 0.13 |
| waves2 | 0.04 | 0.00 | 0.04 | 0.02 | 0.02 | 0.00 | 0.08 | 0.08 | 0.04 | 0.00 | 0.08 | 0.14 | 0.00 | 0.16 | **0.31** |
| **Musical segment** | | | | | | | | | | | | | | | |
| barking3 | **0.24** | 0.05 | 0.00 | 0.00 | 0.05 | 0.08 | 0.05 | 0.11 | 0.14 | 0.14 | 0.05 | 0.00 | 0.05 | 0.03 | 0.00 |
| compr.-stret.3 | 0.00 | **0.33** | 0.00 | 0.06 | 0.06 | 0.00 | 0.06 | 0.11 | 0.00 | 0.00 | 0.00 | 0.00 | 0.28 | 0.06 | 0.06 |
| endl.-traj.3 | 0.00 | 0.00 | **0.35** | 0.05 | 0.08 | 0.03 | 0.00 | 0.13 | 0.00 | 0.03 | 0.08 | 0.08 | 0.08 | 0.05 | 0.08 |
| fading a way 3 | 0.10 | 0.03 | 0.07 | **0.30** | 0.00 | 0.00 | 0.10 | 0.10 | 0.00 | 0.03 | 0.03 | 0.07 | 0.08 | 0.03 | 0.13 |
| falling | 0.13 | 0.09 | 0.16 | 0.00 | **0.22** | 0.00 | 0.06 | 0.03 | 0.06 | 0.00 | 0.13 | 0.00 | 0.06 | 0.06 | 0.00 |
| floating3 | 0.03 | 0.00 | 0.00 | 0.00 | 0.00 | **0.41** | 0.03 | 0.11 | 0.16 | 0.14 | 0.03 | 0.03 | 0.00 | 0.03 | 0.00 |
| in suspense3 | 0.10 | 0.00 | 0.00 | 0.08 | 0.08 | 0.05 | **0.18** | 0.05 | 0.15 | 0.00 | 0.10 | 0.00 | 0.05 | 0.08 | 0.10 |
| moving for.3 | 0.04 | 0.00 | 0.08 | 0.06 | 0.02 | 0.08 | 0.04 | **0.21** | 0.08 | 0.00 | 0.09 | 0.08 | 0.00 | 0.17 | 0.08 |
| obsessive3 | 0.11 | 0.00 | 0.04 | 0.00 | 0.07 | 0.11 | 0.04 | 0.14 | **0.32** | 0.07 | 0.00 | 0.11 | 0.00 | 0.00 | 0.00 |
| propulsion3 | 0.16 | 0.00 | 0.02 | 0.04 | 0.00 | 0.12 | 0.00 | 0.04 | 0.06 | **0.36** | 0.06 | 0.04 | 0.02 | 0.08 | 0.00 |
| spinning3 | 0.00 | 0.00 | 0.06 | 0.06 | 0.03 | 0.00 | 0.13 | 0.16 | 0.00 | 0.03 | **0.29** | 0.10 | 0.00 | 0.03 | 0.10 |
| stationary3 | 0.00 | 0.00 | 0.04 | 0.07 | 0.04 | 0.00 | 0.04 | 0.09 | 0.04 | 0.00 | 0.00 | **0.38** | 0.02 | 0.07 | 0.20 |
| sus.-quest.3 | 0.14 | 0.07 | 0.03 | 0.00 | 0.17 | 0.00 | 0.03 | 0.07 | 0.00 | 0.10 | 0.00 | 0.03 | **0.31** | 0.03 | 0.00 |
| wanting to s.3 | 0.00 | 0.00 | 0.02 | 0.02 | 0.00 | 0.00 | 0.02 | 0.17 | 0.04 | 0.09 | 0.04 | 0.04 | 0.02 | **0.33** | 0.20 |
| waves3 | 0.00 | 0.02 | 0.04 | 0.06 | 0.02 | 0.00 | 0.06 | 0.09 | 0.04 | 0.00 | 0.03 | 0.15 | 0.00 | 0.15 | **0.30** |

**Table 3.** *F* and *p* statistics for the main effect of the Kind of segment (musical segment, segment played on the piano and PTMs), the Expertise of the participants (musician or non-musician) and for the Kind of segment by Expertise interaction, for the association (upper) and confusion (lower) frequencies.

| Association frequency | F | p |
|---|---|---|
| Kind of segment | (2,56) = 0.353 | .70 |
| Expertise | (1,84) = 0.031 | .86 |
| Kind of segment * Expertise | (2,84) = 0.239 | .79 |

| Confusion frequency | F | p |
|---|---|---|
| Kind of segment | (2,56) = 0.295 | .75 |
| Expertise | (1,84) = 3.992 | .04 |
| Kind of segment * Expertise | (2,84) = 0.199 | .88 |

TSU category they came from (grey cells in Table 2). For instance, the segment called fading away2 was grouped 0.22 time with segments from the *Moving forward* TSU and 0.11 time with segments from its TSU.

Moreover, mismatching can occur with a few other TSU categories or with many other categories. Confusion frequency is an erroneous association frequency in the number of relevant TSUs. In other words, it is the number of times a segment was associated with another TSU segment, divided by the number of relevant TSUs. For example, the musical segment compressing–stretching out1 was grouped together 0.10 times with segments from the *Moving forward* TSU, 0.15 times with segments from the *Braking* TSU, 0.10 times with segments from the *Fading away* TSU, 0.15 times with a segment from the *Falling* TSU, 0.05 times with segments from the *Floating* TSU *and* 0.05 time with the *Waves* TSU. Consequently, its confusion frequency was the sum of these frequencies divided by six, which gives 0.10. Thus, if all mistakes concern only one other TSU, the confusion frequency will be more important. A weak or medium association frequency associated with a strong confusion frequency, i.e., with mistakes shared by only one or two TSU(s), indicates that there is confusion between TSUs. A TSU is relevant when its association frequency is strong and when the few mistakes are shared by several TSUs.

By having sound segments as random variables, we run ANOVAs. An ANOVA including the factors Kind of segment (musical excerpts, piano, PTMs) and Expertise (musicians *vs.* non-musicians) were computed on the association and confusion frequencies. Results are presented in Table 3. No significant effect of Kind of segments, Expertise, or interaction between those two factors was observed on the association frequency. On the confusion frequency, no significant effect was found for Kind of segments, but effect of Expertise was significant ($F(1,84) = 3.99$, $p < .05$), with a confusion frequency of 0.14 among musicians and 0.11 among non-musicians. The interaction between the Expertise and the Kind of segments was not significant.

Figures 8A and 8B allow for the comparison of the association and confusion frequencies for all TSUs (taking all kinds of segments together) for both groups (musicians, Figure 8A, and non-musicians, Figure 8B). In the musicians group, the *Propulsion, Waves, Wanting to start, Stationary, Suspending–questioning* and *Endless trajectory* TSUs were well recognized, whereas the *Spinning* TSUs showed a significant confusion frequency (0.48). This *Spinning* TSU was particularly associated with segments from the *Moving forward* TSU by non-musician participants.
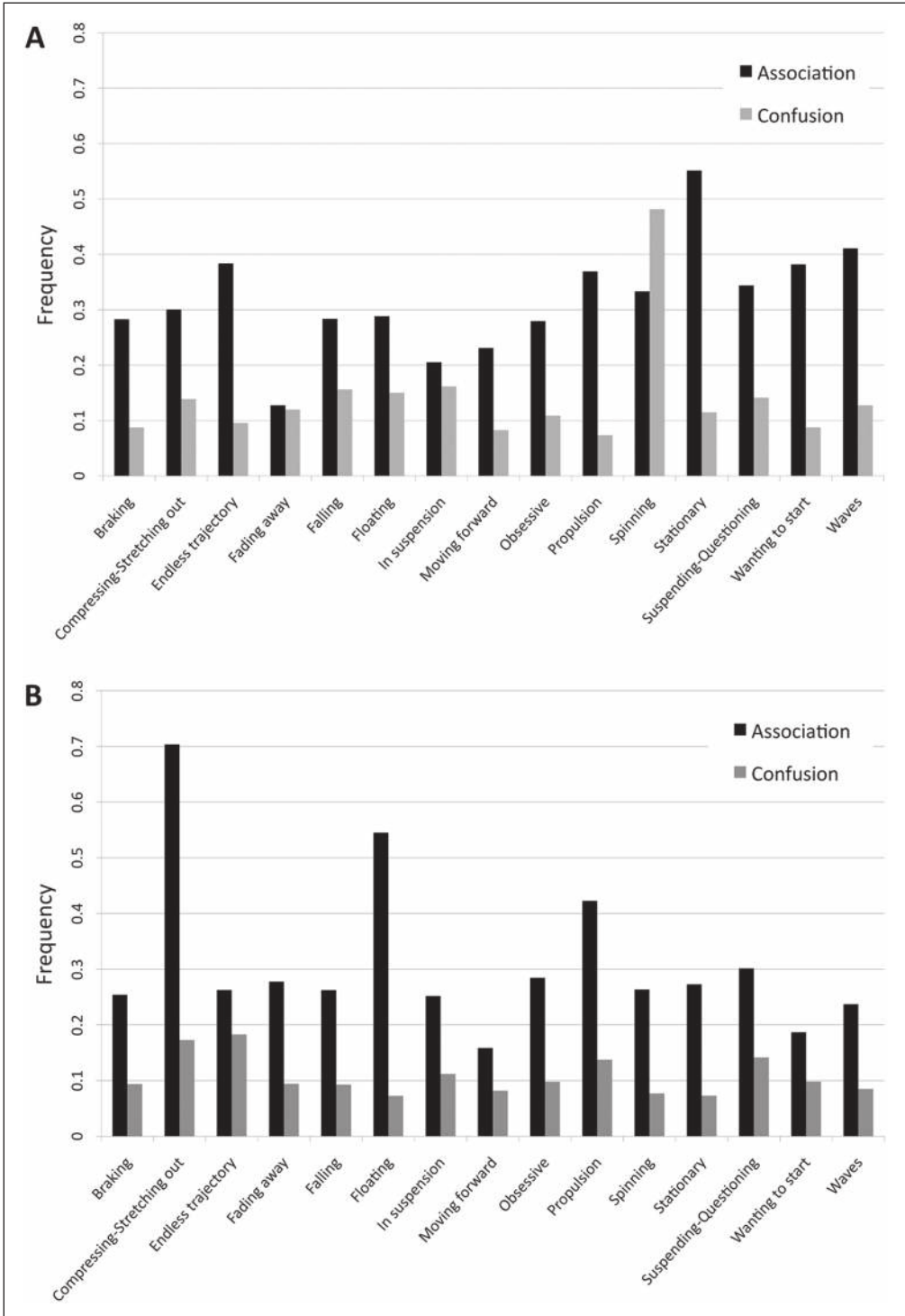
**Figure 8.** Association and confusion frequencies, for each TSU, for musicians (A, upper graph) and for non-musicians (B, lower graph).

## Discussion

The main objective of our experiment was to test the validity of PTMs as a good way of modelling TSUs and as psychological categories of temporal musical processes equivalent to TSUs. Instead of using the spectral envelope or texture of the audio signal (e.g., Grasser, Flexer, & Grill, 2011) to evaluate how much a PTM is similar to its corresponding TSU, we resorted to a categorization task, in which musician and non-musician participants had to group together different kinds of musical segments that corresponded to TSUs. This is a sounder validation method, not only because it is based on human judgment, but because the task was to group musical segments that were meant to be different while belonging to the same category.

Results showed no difference in the grouping performance (i.e., association and confusion frequencies) of musical segments according to their mode of production, namely TSUs taken from existing musical works, TSUs played on the piano and PTMs. PTMs present association and confusion frequencies that are similar to the two other kinds of segments. Thus, PTMs seem to be a conclusive modelling of the TSUs. Let us remember that the only common parameter of these three kinds of segment is their temporal progression. Then, the originality of our results is to show that this common denominator is well perceived by listeners and allows them to recognize a musical form regardless of its format.

As seen in the introduction to the article, TSUs were originally created for electroacoustic music and have been applied to classical music. Once again, our results show that the TSU categories could be applied to different musical formats, and we agree on this point with Delalande et al. (1996, p. 70) for whom "a notable characteristic of the TSU typology consists in not being linked to such a musical construct." This is very relevant for the PTMs, whose vocabulary and concepts are totally independent from the media to which they are applied.

Through the approach that has led to the creation of TSUs and PTMs, the media is the sound, but there are already a few attempts to extend their use to other areas such as painting (Mandelbrojt, 2008) or dancing (Feïlane & Tijus, 2008). Some studies have emphasized the importance of musical soundtracks on the segmentation and communication in movies (Lipscomb & Tolchinsky, 2005; Marshall & Cohen, 1988), and it would be interesting to show in a more systematic way the presence of such temporal units in the visual field and validate the general hypothesis of the existence of temporal process units generic to temporal arts (Bootz & Hautbois, 2010b).

In our experiment, we also wanted to observe a possible effect of the participants' expertise on the perception of TSUs and PTMs. Regarding the number of groups formed, results show that musicians seem to do better than non-musicians, or rather to make a number of groups that is closer to the number of TSU categories. However, the confusion frequency is significantly more important among the musician participants than among the non-musician ones. Our results have not allowed us to conclude that the musicians' expertise had a significant impact.

As is a convention in the research on music perception, we made a distinction between non-musician and musician participants. The latter have learnt to play a music instrument, read a score, understand a musical work, and so forth. Apparently, this distinction is not relevant in our experiment and musical training is not necessary to perceive the TSUs or PTMs. This result is in line with those of Frey, Daquet et al. (2009), and shows that perceptive and cognitive processes involved in the categorization of TSUs would be independent from explicit musical knowledge. Recent data about different aspects of musical perception has also shown that listeners without musical training perceive music as well as trained musicians do (for a review, see Bigand & Poulin-Charronnat, 2006). They also suggest that some kind of implicit learning occurs when one passively listens to music, which allows non-musicians to develop a real

expertise in the field of music (Bigand, 2004), through a mere "exposure" to music, without any help from explicit learning (Bigand, 2006). The role of implicit learning is important given the richness of musical stimulation in everyday life and the remarkable ability of humans to mentally internalize the external regularities of the auditory environment (Saffran, Aslin, & Newport, 1996; Saffran, Johnson, Aslin, & Newport, 1999). Therefore, the perception and categorization of TSUs and PTMs may well rely on knowledge acquired through implicit learning.

The categorization and organization of knowledge rely on environmental regularities in order to favor some categories that are more informative than others. As the environment comes first, i.e., it is not chosen by the individual, it is thus the environment that determines the formation of the categories. Those categories reflect the intrinsic structure of the perceived world (Houix, 2003, p. 45). The temporal process would be a characteristic partaking in the sound environment that individuals would integrate through implicit learning.

Moreover, we can observe that the different TSU categories do not present equivalent association and confusion frequencies. As previously observed in Frey, Daquet et al. (2009), the TSUs *Spinning* and *Moving forward* are often mixed up. The TSU *Floating* was also particularly consistent in the Frey, Daquet et al. (2009) experiment, as well as in this experiment. However, the TSU *Braking* was one of the most consistent in the Frey, Daquet et al. (2009) experiment, whereas it is one whose categorization was the poorest in this experiment. This can be partly explained by the fact that the sound segments presented in both experiments were different, and some of them may be more prototypical than others.

Further experiments could certainly clarify these results, but all the same, the question raised is that of the number of the TSU categories. Indeed, researchers from the MIM have found 19 categories, clearly defined with established boundaries, but why would there not be more categories? Or, on the contrary, some TSU categories could be redundant and grouped together? Another explanation of those differences between categories lies in the status of some extracts used in our experiment. It has been shown that all the elements of a category do not have the same importance; some of them are more representative than others and hold a more salient position (Rosch, 1973; Rosch & Mervis, 1975). Barsalou (1983) emphasized the existence of uncertain cases, that is to say items whose category is unclear. He has shown that subjects do not know whether "a radio" belongs to the category "furniture" or not. Conversely, the element "chair" is more representative of the category "furniture" than the element "armchair," and we can argue that, among the chosen segments used in our experiment, some of them would be more prototypical than others. Another explanation would rely on the intrinsic complexity of some TSU categories that would be cognitively more intricate. By analogy with geometric figures, circles are figures less complex and more easily recognized than triangles (Pinet & Gentaz, 2007). Another source of difficulty could be relative to the structure of the whole categories of TSUs. Some TSUs share common characteristics, and there may be an underlying hierarchy for those categories, just like a square is a particular rectangle.

Finally, it is difficult to compare TSUs with the MIM's identification cards only: one of the objectives of PTMs is to clarify this point, by providing more precise defining and classificatory criteria. It would now be interesting to determine how close or far these categories of TSU are. This issue is the focus of a work in progress, for which the formalization in PTM and the results of our experiments of categorization can provide possible answers, especially with the future use of additive trees (ADDTREE program, Sattath & Tversky, 1977).

In conclusion, the present study demonstrates the relevance of PTMs in the modelling of TSUs, and constitutes another form of validation of the TSU categories. Our experiment allows us to consider TSUs as perceptive units that make sense for listeners, whether they are

musicians or not. Thus, the TSU as a physical unit leads to the TSU as a cognitive unit. However, if the formalization of the TSU seems to be justified from a theoretical perspective, our results show that the validity of some TSU categories remains questionable. The theorization work on TSUs is still ongoing and could be facilitated by this kind of experiment.

## Acknowledgements

## Funding

## Notes

1. FRIDA (Formalisme pour la Représentation des Intéractions avec un Dispositif Automatique) is a programming environment written in Java and dedicated to cognitive psychology experiments, which allows for the presentation of various stimuli and recording of responses. The FRIDA software includes: (1) a graphic editor; (2) an automata editor; and (3) a runtime process.
2. See the introduction to this article. The term "acousmatic" was first used by Pierre Schaeffer and refers to sounds that one hears while unaware of the underlying causes. In an acousmatic listening, the attention shifts away from the physical object that induces auditory perception back towards the content of the perception.

## References

Alain, C., Arnott, S. R., & Picton, T. W. (2001). Bottom-up and top-down influences on auditory scene analysis: Evidence from event-related brain potentials. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(5), 1072–1089.

Austin, M. (2011). A case study in analyzing musical multimedia using Unités Sémiotiques Temporelles. *New Sound*, *38*, 87–95.

Baddeley, A. (1986). *Working memory*. Oxford: Oxford University Press.

Barsalou, L. W. (1983). Ad hoc categories. *Memory and Cognition*, *11*(3), 211–227.

Bayle, F. (2003). *L'image de son: Technique de mon écoute*. In I. Misch & C. v. Blumröder (Eds.), *François Bayle – L'image de son/Klangbilder: Technique de mon écoute/Technik meines Hörens* (Komposition und Musikwissenschaft im Dialog IV – 2000–2003). Münster: LIT.

Berz, W. L. (1995). Working memory in music: A theoretical model. *Music Perception*, *12*(3), 353–364.

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, *94*, 115–147.

Bigand, E. (2004). Musical and non musical brains: A trompe-l'oeil debate? *Neuropsy News*, *3*(2), 68–72.

Bigand, E. (2006). Musiciens et non-musiciens perçoivent-ils la musique différemment? In B. Lechevalier, F. Eustache, & H. Platel (Eds.), *Le cerveau musicien* (pp. 207–236). Bruxelles: De Boeck.

Bigand, E., & Poulin-Charronnat, B. (2006). Are we experienced listeners? A review of the musical capacities that do not depend on formal musical training. *Cognition*, *100*, 100–130.

Bootz, P., & Hautbois, X. (2007). Times measures in documents: The model of "Motifs Temporels Paramétrés." In R. Skare, N. W. Lund, & A. Vårheim (Eds.), *A document (re)turn* (pp. 197–222). Frankfurt am Main: Peter Lang. Retrieved from http://halshs.archives-ouvertes.fr/SHS/sic_00637789/fr/

Bootz, P., & Hautbois, X. (2008). Les Motifs Temporels Paramétrés. In E. Rix, & M. Formosa (Eds.), *Vers une sémiotique générale du temps dans les arts. Actes du colloque "Les Unités Sémiotiques Temporelles (UST), nouvel outil d'analyse musicale: Théories et applications"* (pp. 139–176). Sampzon, France: Delatour.

Bootz, P., & Hautbois, X. (2010a). Modélisation des structures temporelles par les MTP. *Musimédiane, N°5, Les Unités Sémiotiques Temporelles: Enjeux pour l'analyse et la recherche*. Retrieved from http://www.musimediane.com/spip.php?article107

Bootz, P., & Hautbois, X. (2010b). Les UST visuelles: analyse et création. *Musimédiane, N°5, Les unités sémiotiques temporelles: Enjeux pour l'analyse et la recherche*. Retrieved from http://www.musimediane.com/spip.php?article113

Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound*. Cambridge, MA: MIT Press.

Cohen-Levinas, D. (2005). *Musique et philosophie*. Paris: L'Harmattan.

Couprie, P. (2013) EAnalysis. Retrieved from http://logiciels.pierrecouprie.fr/?page_id=402

Delalande, F. (1990). *Eléments d'analyse de la stratégie de composition*. Marseille: MIM.

Delalande, F. (2008). UST et analyse: Introduction. In E. Rix, & M. Formosa (Eds.), *Vers une sémiotique générale du temps dans les arts. Actes du colloque "Les Unités Sémiotiques Temporelles (UST), nouvel outil d'analyse musicale: théories et applications"* (pp. 17–18). Sampzon, France: Delatour.

Dehaene, S., Izard, V., Pica, P., & Spelke, E. (2006). Core knowledge of geometry in an Amazonian indigene group. *Science*, *311*, 381–384.

Delalande, F., Formosa, M., Frémiot, M., Gobin, P., Malbosc, P., Mandelbrojt, J., et al. (1996). *Les Unités Sémiotiques Temporelles—éléments nouveaux d'analyse musicale* [book-audio CD]. Marseille: MIM/ Documents Musurgia.

Deutsch, D. (1970). Tones and numbers: Specificity of interference in immediate memory. *Science*, *168*, 1604–1605.

Deutsch, D. (1972a). Effect of repetition of standard and comparison tones on recognition memory for pitch. *Journal of Experimental Psychology*, *26*, 229–234.

Deutsch, D. (1972b). Mapping of interactions in the pitch memory store. *Science*, *175*, 1020–1022.

Deutsch, D. (1973). Interference in memory between tones adjacent in the musical scale. *Journal of Experimental Psychology*, *100*, 228–231.

Di Santo, J. L. (2008). Composer avec les UST. In E. Rix, & M. Formosa (Eds.), *Vers une sémiotique générale du temps dans les arts. Actes du Colloque "Les Unités Sémiotique Temporelles (UST), nouvel outil d'analyse musicale, théories et applications"* (pp. 257–270). Sampzon, France: Delatour.

Dubois, D. (1991). *Sémantique et cognition. Catégories, prototypes, typicalité*. Science du langage. Paris: CNRS Edition.

Favory, D., Formosa, M., Frémiot, M., Gobin, P., Malbosc, P., Mandelbrojt, J., et al. (2002). *Les Unités Sémiotiques Temporelles - Nouvelles Clés pour l'Ecoute*. [CD-ROM]. Marseille: Edition MIM.

Feïlane, D., & Tijus, C. (2008). Classification condensée des formes de mouvements communes aux disciplines de danse et leurs rapports aux UST. In E. Rix, & M. Formosa (Eds.), *Vers une sémiotique générale du temps dans les arts. Actes du colloque "Les Unités Sémiotiques Temporelles (UST), nouvel outil d'analyse musicale: théories et applications"* (pp. 301–304). Sampzon, France: Delatour.

Francès, R. (1958). *La perception de la musique*. Paris: Vrin.

Frey, A., Daquet, A., Poitrenaud, S., Tijus, C., Frémiot, M., Formosa, M., et al. (2009). Pertinence cognitive des Unités Sémiotiques Temporelles. *Musicae Scientiae*, *13*(2), 415–440.

Frey, A., Marie, C., Prod'Homme, L., Timsit-Berthier, M., Schön, D., & Besson, M. (2009). Temporal semiotic units as minimal meaningful units in music? An electrophysiological approach. *Music Perception*, *26*(3), 247–256.

Grabócz, M. (2008). Quelques processus archétypiques – ou unités sémiotiques temporelles – dans les écrits et les œuvres de compositeurs contemporains. In E. Rix et M. Formosa (Eds.), *Vers une sémiotique générale du temps dans les arts. Actes du colloque "Les Unités Sémiotiques Temporelles (UST), nouvel outil d'analyse musicale: théories et applications"* (pp. 93–112). Sampzon, France: Delatour.

Grasser, M., Flexer, A., & Grill, T. (2011). On computing morphological similarity of audio signals. In Proceedings of the 8th Sound and Music Computing Conference, 6–9 July 2011, Padova, Italy (pp. 279–282). Padova: Padova University Press.

Hoffmann, D. D., & Richards, W. A. (1984). Parts of recognition. *Cognition*, *18*, 65–96.

Houix, O. (2003). *Catégorisation auditive des sources sonores*. Unpublished doctoral dissertation. University of Maine.

Huron, D. (1991). Auditory scene analysis: The perceptual organization of sound by Albert S. Bregman. *Psychology of Music*, *19*(1), 77–82.

Imberty, M. (1981). *Les écritures du temps: sémantique psychologique de la musique, tome 2*. Paris: Dunos.

Klinkenberg, J-M. (1996). *Précis de sémiotique générale*. Liège : De Boeck &Larcier.

Kramer, J. (2004). Le temps musical. In J.-J. Nattiez (Ed.), *Musiques—Une encyclopédie pour le XXIe sciècle, 2, Les savoir musicaux* (pp. 189–213). Arlos: Actes Sud / Cité de la Musique.

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Science*, *4*, 463–470.

Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature*, *307*, 161–163.

Landy, L. (2007). *Understanding the art of sound organization*. Cambridge, MA: The MIT Press.

Lipscomb, S. D., & Tolchinsky, D. E. (2005). The role of music communication in cinema. In D. Miell, R. MacDonald, & D. Hargreaves (Eds.), *Musical Communication* (pp. 383–404). Oxford: Oxford University Press.

Mâche, F.-B. (2001). *Musique au singulier*. Paris: Odile Jacob.

Mandelbrojt, J. (2008). Arts plastiques et UST. In E. Rix, & M. Formosa (Eds.), *Vers une sémiotique générale du temps dans les arts. Actes du colloque "Les Unités Sémiotiques Temporelles (UST), nouvel outil d'analyse musicale: théories et applications"* (pp. 313–316). Sampzon: Delatour-France.

Marshall, S. K., & Cohen, A. J. (1988). Effects of musical soundtracks on attitudes toward animated geometric figures. *Music Perception*, *6*(1), 95–112.

Miereanu, C. (1996). *Fuite et conquête du champ musical*. Paris: Meridiens Klincksieck.

Milner, P. M. (1974). A model for visual shape recognition. *Psychological Review*, *81*, 521–535.

Parmegiani, B. (2001). *De natura sonorum* (LP, 1976; reissued on CD, 2001, C3001). Paris: INA-GRM.

Pechman, T., & Mohr, G. (1992). Interference in memory for tonal pitch: Implications for a working-memory model. *Memory and Cognition*, *20*, 314–320.

Peretz, I., & Zatorre, R. J. (2005). Brain organization for music processing. *Annual Review of Psychology*, *56*, 89–114.

Petitot, J. (1989). Perception, cognition et objectivité morphologique. In S. McAdams, & I. Deliège (Eds.), *La musique et les sciences cognitives* (pp. 243–256). Liège, Bruxelles: Pierre Mardaga.

Pinet, L., & Gentaz, E. (2007). La reconnaissance des figures géométriques planes par les enfants de 5 ans. *Grand N*, *80*, 17–28.

Risset, J.-C. (2001). *Mutations* [LP, 1970; reissued on CD, C1003]. Paris: INA-GRM.

Rosch, E. H. (1973). Natural categories. *Cognitive Psychology*, *4*, 328–350.

Rosch, E. H., & Mervis, C. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, *7*, 573–605.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month old infants. *Science*, *274*, 1926–1928.

Saffran, J. R, Johnson, E. K, Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tones sequences by human infants and adults. *Cognition*, *70*(1), 27–52.

Sattath, S., & Tversky, A. (1977). Additive similarity trees. *Psychometrika*, *42*, 319–345.

Schaeffer, P. (1966). *Traité des objets musicaux*. Paris: Editions du Seuil.

Sciarrino, S. (1998). *Le figure della musica da Beethoven a oggi*. Milano: Ricordi.

Smalley, D. (1986). Spectro-morphology and structuring processes. In S. Emmerson (Ed.), *The language of electroacoustic music* (pp. 61–96). London: Macmillan.

Tarr, M. J. (1995). Rotating objects to recognize them: A case study of the role of viewpoint dependency in the recognition of three-dimensional objects. *Psychonomic Bulletin and Review*, *2*, 55–82.

Tervaniemi, M., Just, V., Koelsch, S., Widman, A., & Schröger, E. (2005). Pitch discrimination accuracy in musicians vs nonmusicians: An event-related potential and behavioral study. *Experimental Brain Research*, *161*, 1–10.

Thibaut, J. P. (1997). Similarité et catégorisation. *L'Année psychologique*, *97*(4), 701–736.

Tijus, C., & Cordier, F. (2003). Psychologie de la connaissance des objets. Catégories et propriétés, tâches et domaines d'investigation. *L'année psychologique*, *103*, 223–256.

Williamon, A., & Egner, T. (2004). Memory structures for encoding and retrieving a piece of music: An ERP investigation. *Cognitive Brain Research*, *22*, 36–44.

Williamon, A., & Valentine, E. (2002). The role of retrieval structures in memorizing music. *Cognitive Psychology*, *44*, 1–32.

Wishart, T. (1994). *Audible design: A plain and easy introduction to practical sound composition*. York, UK: Orpheus The Pantomime.

## Appendix

### Invariant TSUs

#### Invariant TSUs by repetition
#### Waves

**Morphological description**: One phase unit, non-time-delimited, made up of the slow repetition of an increasing then decreasing sound motif. The shape of the profile can concern different morphological criteria (mass, dynamics, grain, etc.).

**Semantic description**: Each cycle conveys the feeling of being pushed forward and then driven back until the end. We get the impression that we are stagnating through this unit although we feel motion inside each cycle.

#### Spinning

**Morphological description**: One-phase unit, non-time-delimited, in which a parameter (pitch, timbre) is driven by a quick cyclic repetition, along with a thrust in each cycle.

**Semantic description**: We have the feeling that an object is spinning around itself or in space.

#### Obsessive

**Morphological description**: One-phase unit, non-time-delimited, with a quick and possibly varied repetition of a pulsed element.

**Semantic description**: We feel constrained by a mechanical process on which we cannot seem to act.

#### Invariant TSUs by stagnation
#### Stationary

**Morphological description**: Non-time-delimited unit with a slow temporal evolution. It shows a global temporal regularity but random elements may appear on another level. Several morphological configurations can be found: scarcely changing temporal structure, random details in global permanency or slow cycles with little variety.

**Semantic description**: The unit seems to get nowhere: there is no progress even if there is always something happening.

#### Floating

**Morphological description**: Non-time-delimited unit with a slow temporal evolution. Irregular succession of disjointed and relatively brief temporal elements.

**Semantic description**: Despite the sound events that randomly appear, there is no feeling of expectation or suspense. The way the objects are distributed in time is perceived as a linear flow.

### In suspension

**Morphological description**: Non-time-delimited unit made up of a repeated sequence, without significant variation, with a slow temporal progression. The sound matter and temporal events do not change much.

**Semantic description**: The balance of the operating forces conveys a feeling of stillness related to a sense of hesitant expectation. We know that something will happen, but we do not know when or what.

### Invariant TSUs by chaotic effect
### Divergent

**Morphological description**: Non-time-delimited unit characterized by a succession of brief temporal sequences without any relation to one another.

**Semantic description**: Lack of causality between successive temporal events. The great number of directions, without apparent connections, results in indecision.

### Chaotic

**Morphological description**: Non-time-delimited unit characterized by a dense superimposition of temporal sequences without any relation to one another.

**Semantic description**: Lack of causality between superimposed temporal events. It causes an indescribable hubbub, a high saturation, a feeling of incommunicability and tension.

## Variant TSUs

### Variant TSUs with a uniform development
### Moving forward

**Morphological description**: One-phase unit, non-time-delimited, with a globally uniform and uninterrupted process in which a brief repeated sequence regularly renews the driving activity.

**Semantic description**: It seems to regularly push us forward. It gives us the impression that we are moving decisively in a precise direction.

### Endless trajectory

**Morphological description**: One-phase unit, non-time-delimited, with a linear and usually slow evolution of a sound parameter.

**Semantic description**: The process must be oriented in a direction (for example, upwards or downwards) and yet it seems to never end. The sound phenomenon must be long enough to be perceived as a process and not an ephemeral event.

### Heaviness

**Morphological description**: Non-time-delimited unit characterized by a brief sequence slowly and irregularly repeated.

**Semantic description**: Seems to have trouble progressing, despite a driving energy.

*Variant TSUs with a thwarted development*
*Fading away*

**Morphological description**: Time-delimited unit with a single phase. Progressive extinction of resonance type, without any complementary energy supply.

**Semantic description**: It is like the trajectory of a sailing boat with its sails down, which moves on, slowing down, because of momentum. Predictable process until its conclusion.

*Braking*

**Morphological description**: Time-delimited unit with two phases: a process that has already started is suddenly opposed to another process with a decreasing energy.

**Semantic description**: It gives the impression of a forced slowdown. All movement is abruptly refrained until it stops.

*Stretching*

**Morphological description**: Single-phase time-delimited unit. Linear growth of a morphological trait, with increased energy, on a frozen temporal structure.

**Semantic description**: It seems to go forward to the maximum of a process (or an effort). Feeling of a stretching of sound involving two opposite forces.

*Wanting to start*

**Morphological description**: Non-time-delimited unit made up of two repeated phases. The first is an articulate shape, relatively short. The second is the opposite of the first (mass, dynamics, homogeneity). The reiteration of the two phases is not strict: there is at least one parameter that varies.

**Semantic description**: It is like something trying to start. Reiteration suggests a repeated effort to reach a goal.

*Variant TSUs with a disrupted balance*
*Falling*

**Morphological description**: Two-phase time-delimited unit. The first phase is globally uniform while the second is an accelerated change in pitch (upwards or downwards). The limit of this transition is sharp, not continuous.

**Semantic description**: A breaking unstable state of equilibrium. Suspension followed by pivoting (the suspension phase is realized after the fact).

*Propulsion*

**Morphological description**: Three-phase time-delimited unit. The first phase is quite a sustained fulcrum: a prolonged and homogeneous sound or slow iteration, globally uniform. The following phase is a brief acceleration of intensity, pitch or any other morphological trait. The third phase is a typical resonance or silence.

**Semantic description**: We feel the application of a force to a steady state, resulting in an accelerated movement. Projection from a starting point.

*Compressing–Stretching out*

**Morphological description**: Time-delimited unit with two contrasting phases. During the "compressing" phase, the sound matter is discontinuous and erratic. The "stretching out" phase is a globally uniform segment.

**Semantic description**: First, there is a feeling of compression (as if we pressed down hard on an obstacle), then the barrier is suddenly overcome, suppressing all resistance and releasing the power. It is a sudden shift from localized energy to scattered energy.

*Suspending–Questioning*

**Morphological description**: Time-delimited unit with two contrasting phases. The first phase is a brief and usually varied and repeated temporal sequence. The second phase morphologically contrasts with the one: a short sustained and rising sound or a silence.

**Semantic description**: An interrupted movement in a frozen position.