

Temporal Semiotic Units as minimal meaningful units in music? An electrophysiological approach.

Aline Frey, Céline Marie, Lucie Prod'Homme, Martine Timsit-Berthier, Daniele Schön, Mireille R Besson

▶ **To cite this version:**

Aline Frey, Céline Marie, Lucie Prod'Homme, Martine Timsit-Berthier, Daniele Schön, et al.. Temporal Semiotic Units as minimal meaningful units in music? An electrophysiological approach.. Music Perception, 2009, 26, pp.247 - 256. 10.1525/MP.2009.26.3.247 . hal-02430621

HAL Id: hal-02430621
https://hal.science/hal-02430621

Submitted on 7 Jan 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Music2603_08 1/13/09 12:05 PM Page 247

Temporal Semiotic Units as Minimal Meaningful Units in Music? An Electrophysiological Approach

Aline Frey
Université Paris 8, Saint-Denis, France

Céline Marie
CNRS—Aix-Marseille Université, Marseille, France

Lucie Prod'Homme and Martine Timsit-Berthier
Laboratoire Musique et Informatique de Marseille, Marseille, France

Daniele Schön and Mireille Besson
CNRS—Aix-Marseille Université, Marseille, France

THE AIM OF THIS STUDY WAS TO DETERMINE WHETHER conceptual priming occurs between successively presented short musical pieces called Temporal Semantic Units (TSUs). Behavioral and ERP data were recorded while participants, experts and nonexperts in TSUs, were listening to pairs of TSUs and were asked to determine whether the target TSU evoked the same or a different concept than the prime TSU. Target TSUs were either congruous (i.e., they developed the same musical concept as the prime TSUs) or incongruous (i.e., they started as congruous TSUs but shifted midstream into a different concept). Results showed that, whereas P3a components were elicited in both groups by the shifting into incongruous TSUs, thereby reflecting an automatic shift of attention when the changes occurred, P3b components were elicited in experts and N400-like components were found in nonexperts. The functional significance of these results is discussed in regard of previous results with environmental sounds.

Received May 29, 2008, accepted October 25, 2008.

Key words: event-related potentials (ERPs); N400; temporal semantic units (TSUs); musical semantics; expertise

WHAT IS THE NATURE OF MUSIC? What does music express? What effects does music have on listeners? These are questions that remain unsolved (Daniélou, 1993). When people are asked to describe what they hear while listening to music, their answers are very different, but people with or without formal music education, are most often able to feel the expressive aspects of music (Bicknell, 2002). Uncovering the foundations of the expressive power of music is a central problem in music research. Several theories have been proposed to address this issue. They have focused on the self referential nature of music (Meyer, 1956) or on its temporal/harmonic structure (Jones, 1981, 1982; Lerdahl & Jackendoff, 1983). Note, however, that all these theories have been elaborated within the framework of tonal music, which is to be expected since music is most often based on tonal hierarchies (Dowling & Harwood, 1986). Interestingly, however, the last decades have seen the development of new styles of music within different frameworks (e.g., atonal, serial). New approaches and tools are therefore required for their analysis. McAdams (1989) has already emphasized the need to leave the tonal system, even if "psychologists benefit from a larger experience in the system than outside" (p. 16). As more and more music relies on other systems, other parameters have to be considered to understand musical processing.

TSUs as Minimal Segments of Meaning

A starting point for the analysis of electroacoustic music was the work of Pierre Schaeffer. In the "Traité des Objets Musicaux" (1966), Schaeffer proposed a descriptive vocabulary and a classification grid for musical objects. This description relied on reduced listening and implied abstraction of the sound source (i.e., acousmatic listening). However, in his research, Schaeffer did not consider the temporal and evolving aspects of musical objects. And music is mainly an art of time. As Francès notes (1958), "The experience of the listener is rigorously irreversible and successive" (p. 202).

Music Perception volume 26, issue 3, pp. 247–256, issn 0730-7829, electronic issn 1533-8312 © 2009 by the regents of the university of california. all rights reserved. please direct all requests for permission to photocopy or reproduce article content through the university of california press's rights and permissions website, http://www.ucpressjournals.com/reprintinfo.asp. DOI:10.1525/mp.2009.26.3.247

Music2603_08 1/13/09 12:05 PM Page 248

Sounds exist only in and through time, with consciousness and memorization. These considerations led researchers and composers of the Laboratoire Musique et Informatique de Marseille (MIM) to develop a new tool of musical analysis intended for electroacoustic music and based upon music temporal structure.[1] They developed the concept of Temporal Semiotic Units (TSUs), following which musical segments convey meaning through their dynamic organization in time. To define the different TSU categories, they used a methodology based on collective listening of many different musical pieces. During the listening sessions, these researchers attempted to identify the specific temporal forms that developed between two and several seconds and that would convey meaning. A total of 19 TSU categories were thereby identified and named using representative labels: Chute, Trajectoire inexorable, Contracté-étendu, Elan, Etirement, En Flottement, Sans direction par divergence d'information, Lourdeur, Freinage, Obsessionnel, Qui avance, Qui tourne, Qui veut démarrer, Sans direction par excès d'information, Suspension- interrogation, En suspension, Par vagues, Stationnaire and Sur l'erre.[2] Finally, a written description of each TSU was also established that included a global morphological description, a semantic description, and other relevant characteristics, such as energy, intensity, and tension.

N400 as an Electrophysiological Index of Semantics

Event-related potentials (ERPs) are derived from the electroencephalogram (EEG). They reflect the changes in brain electrical activity associated with a specific event. Different components can be defined by their polarity, amplitude, latency, and scalp distribution, and are considered to reflect specific perceptual and cognitive processes. Kutas & Hillyard (1980) first described the ERP component N400, characterized as a negative deflection distributed over central-parietal sites and peaking approximately 400ms after stimulus onset. Many studies have shown that this component is enhanced for semantically incongruous sentence-ending words relative to congruous ones, the so-called "N400 effect." For example, in the sentence: "The zebras ran away, chased by . . . " the final word "rabbits" is associated with larger N400 component than "lions." There is a consensus that the N400 is a good index of the lack of integration of a word within its preceding semantic context (Kutas & Federmeier, 2000), and that it may reflect the cognitive processes associated with semantic expectancy (Kutas & Hillyard, 1984).

Further investigations have shown that N400-like components also are observed with nonlinguistic stimuli such as objects (Barrett & Rugg, 1990), faces (Barrett & Rugg, 1989; Chaby, Jemel, George, Renault, & Fiori, 2001), music (Koelsch et al., 2004) or environmental sounds (Cummings et al., 2006; Orgs, Lange, Dombrowski, & Heil, 2006, 2007; Plante, Van Petten, & Senkfor, 2000; Van Petten & Rheinfelder, 1995) when they do not semantically or conceptually match the preceding context. Thus, the N400 has been found to be modulated by priming; that is, the facilitatory effect of having previously encountered a related stimulus on its subsequent processing (Koyama, Nageishi, & Shimok-Chi, 1992). However, to our knowledge, all auditory priming experiments constructed so far have always included words either as primes or targets. For instance, Koelsch et al. (2004) have recently investigated semantic priming of words using sentences and music as primes. They reasoned that if music conveys specific meaning, some musical excerpts should prime a given word more than another. Thus, sentences and musical primes were either related or unrelated, in meaning or expressive power, to a target word. As expected, results showed that word targets presented after semantically unrelated sentences elicited N400 components. Interestingly, an effect very similar in latency, scalp distribution, and amplitude was found when semantically unrelated musical excerpts preceded target words. These results were interpreted as showing that "music . . . can prime representations of meaningful concepts" (p. 306) and that "music transfers considerably more semantic information than previously believed" (p. 306). However, it first should be noted, that the N400 component was elicited by a word rather than by the music itself. Second, and as acknowledged by the authors, the duration of the musical excerpts was long enough for participants to mentally generate, or to anticipate, a word or a set of words that matched with music. The question of whether a musical excerpt can prime another musical excerpt remains open.

[1]Temporal dynamics have been described in six large families by Bootz and Hautbois (2007a, 2007b): sound parameters remain constant, evolve monotonously over time, evolve exponentially, first grow and then decrease, follow a Dirac function, or have a specific form.

[2]These labels can be translated as: Falling, Inexorable trajectory, Contracted-Extended, Enthusiastic, Stretching, Floating, Without direction due to divergence of information, Heaviness, Braking, Obsessive, Advancing, Turning, Trying to start up, Without direction due to excess of information, Suspension-Questioning, In suspension, In waves, Stationary, and Coasting.

Music2603_08 1/13/09 12:05 PM Page 249

The first aim of the present experiment was to determine whether a conceptual change occurring within a TSU (i.e., a TSU expressing "falling" shifting in midstream into another TSU expressing "stationary"[3]) would generate an N400 effect. In order to establish a context and activate a specific TSU category, a congruous TSU always was presented first (prime), followed by a congruous TSU from the same category or by an incongruous TSU (target). We hypothesized that if the violation of the concept conveyed by the TSU is processed as a violation of meaning, an N400 component should be generated. However, if the transformation of the TSU is processed at a perceptual (rather than cognitive) level, P300 components may be generated. Indeed, extensive research has shown that unexpected and rare events at the perceptual level elicit P3a components with a fronto-central scalp distribution (Sutton, Braren, Zubin, & John, 1965). The occurrence of P3a components is taken to reflect an involuntary orientation of attention toward the surprising event (Courchesne, Hillyard, & Galambos, 1975) and is generally followed by a P3b component, signaling the conscious registration of the unexpected event (Donchin, 1981).

The Effects of Expertise

The second aim of the present experiment was to test for the effect of musical expertise. Analysis of behavioral data typically shows that the level of performance of musicians is higher than nonmusicians in explicit (Koelsch, Schröger, & Tervaniemi, 1999; Tervaniemi, Just, Koelsch, Widman, & Schröger, 2005) but not implicit music tasks (e.g., Bigand & Poulin-Charronnat, 2006). The latter results highlight the fact that some musical capacities are acquired through mere exposure to music without the help of explicit training. Thus, in some cases, nonmusicians respond to music in the same way as experienced musicians. The role of implicit learning is important given both the richness of musical stimulation in everyday life and the remarkable ability of humans to internalize regularities of the auditory environment (Saffran, Aslin, & Newport, 1996; Saffran, Johnson, Aslin, & Newport, 1999). In contrast, results of experiments using brain imaging methods such as MRI, fMRI, TEP or ERPs typically have shown differences in the anatomofunctional organization of musicians' and nonmusicians' brains (Seung, Kyong, Woo, Lee, & Lee, 2005). For example, Gaser and Schlaug (2003) have shown gray matter volume differences in motor, auditory, and visuospatial brain regions between professional musicians and a matched group of amateur musicians and nonmusicians.

Because we used an explicit categorization task that required explicit knowledge of the concepts conveyed by the TSUs, we hypothesized that TSU experts should recognize congruent TSUs better than nonexperts. Moreover, our main interest was to determine whether TSU experts also would detect incongruous TSUs better than nonexperts, and whether such potential differences would be reflected in the pattern of brain waves.

Method

Participants

Twenty-four participants (mean age = 32.74 years; *SD* = 10.46; eight females, sixteen males; two left-handed) without hearing or neurological disorders participated in this study. Twelve were nonmusicians and twelve were musicians who had at least 14 years of formal music training and were actively engaged in electroacoustic music classes. All participants were paid for their participation. Two participants in the nonexpert group were excluded from further analyses because of a large number of artifacts in the electrophysiological data (blinks and amplifier saturation linked to movements).

Musicians were TSU experts recruited from two classes of musical formation and electroacoustic composition classes at the Laboratoire Musique et Informatique de Marseille. In these classes they learned the concept of TSUs as well as how to use TSUs for their musical compositions. We conducted two pretests to assess their expertise. In both tests, they were asked to identify recorded TSUs that were sung by a professional teacher (L. Prod'Homme). One pretest was a forced-choice test in which participants chose between two possible answers (e.g., Obsessive/Inexorable trajectory). Mean correct responses was 18.40 out of 20 (SD = 1.80). The other pretest was a free-choice test in which participants wrote the names of the TSUs they heard. Mean correct responses were 15.20 out of 20 (*SD* = 3.25). Two participants who performed far below average (scores of 8/20 and 9/20 on the second pretest) were excluded from the study. Note that the experts had never heard the specific TSUs prior to the experiment.

Materials

Of the 19 categories of TSUs first described by Delalande et al. (1996), 12 were used in the experiment. The seven other TSU categories were not considered

[3]See sound example at http://www.incm.cnrs-mrs.fr/pperso/attach/schon/TSUs_material.zip

Music2603_08 1/13/09 12:05 PM Page 250

because their exemplars were too long in duration. A composer from the MIM (L. Prod'Homme) composed new exemplars of the 12 TSUs categories retained for the experiment by strictly following up the rules of construction as defined by Delalande et al. (1996).

Thus, a total of 90 new musical excerpts were composed and played on the piano. Sixty excerpts were constructed to convey a specific concept: congruous TSUs, with Contracted-Extensive (6 exemplars), Enthusiastic (9 exemplars), Stretching (6 exemplars), Floating (12 exemplars), Obsessive (9 exemplars), In waves (9 exemplars), Trying to start up (3 exemplars), Turning (3 exemplars), Without direction due to excess of information (3 exemplars), Stationary (3 exemplars), Suspension-Questioning (9 exemplars), and Inexorable Trajectory (18 exemplars). Thirty incongruous TSUs also were composed: they started as one TSU category but, after a few seconds, they shifted in midstream to another TSU category. The composer was careful that such a transformation did not involve any sudden differences in intensity or frequency. Five TSU experts who did not participate in the study validated the fact that these TSUs were incongruous (i.e., they were able to notice a change of concept within the TSU with no abrupt changes in frequency or intensity).

We used a priming design in which the prime TSUs that were always congruous (30; (mean duration = 10.59 s, *SD* = 5.02 s) were presented twice, once followed by a different congruous target TSU from the same category (30; mean duration = 11.28 s, SD = 4.91 s) and once followed by an incongruous target TSU (30; mean duration = 11.24 s, *SD* = 5.18 s). Within each pair, the target TSU followed the prime TSU after a 2.5-s interval.

Procedure

Each participant listened to 60 TSU pairs that were presented in three blocks of 20 trials. In each block, pairs were presented in a random order, and the order of presentation of the blocks was counterbalanced across participants. Participants were informed that they were to be presented with pairs of musical excerpts (primes and targets) and were asked to press one response button at the end of the second excerpt (target) if they thought that it evoked the same concept as the first, and another button if they thought that the concept evolved to be different from that of the first excerpt (prime). Note that incongruous target TSUs always started with the same concept as the prime but shifted midstream into another concept (i.e., into a different TSU category). There was no time pressure. The side (right or left hand) of the response was counterbalanced across participants. After the end of the target TSU, a row showing XXXX appeared for 2 seconds on the computer screen in front of the participants, and they were asked to blink during this time interval to avoid blinking during data collection.

The experimental session was preceded by a training block, during which participants listened to six pairs of TSUs (three pairs congruous and three pairs incongruous) that were not used in the experimental session.

Data Acquisition

For incongruous excerpts, ERP recordings were time-locked to the moment when the TSU started to be incongruous, as indicated by the composer and as validated by several TSU experts who did not participate to the study. In order to synchronize ERP recordings in congruous TSUs, we computed the average ratio between the entire duration of all incongruous targets and the averaged time of occurrence of the incongruity, and we set the triggers for the congruous item. This ratio was equal to .61 (half duration of the TSUs is .50).

EEG was continuously recorded and then segmented into EEG epochs of 3200 ms, starting 200 ms before the onset of the incongruity for incongruous TSUs or the equivalent time point for congruous TSUs. EEGs were recorded from 32 scalp electrodes located at standard left and right hemisphere positions, and was rereferenced offline to the algebraic average of the right and left mastoids. Both the horizontal and vertical electrooculograms (EOG) were recorded from electrodes located over the external canthi and below the right eye, respectively. Trials containing ocular artifacts, movement's artifacts, or amplifier saturation were excluded from the averaged ERP waveforms.

ERPs were averaged for correct responses only and analyzed in selected latency windows defined both according to the literature (e.g., the N400 component typically develops between 300 ms and 600 ms post-stimulus onset) and visual inspection of the ERP traces: 0-150 ms, 150-300 ms, 300-600ms, and 600-1000ms.

Data were subjected to ANOVAs and *p* values are reported after the Greenhouse-Geisser correction for nonsphericity. Tukey tests were used for posthoc comparisons. Expertise (Experts vs. Nonexperts) was the between-subject factor. The within-subject factors were Electrodes (Fz, Cz, Pz, Oz) for midline analyses, Hemisphere (HG vs. HD), Anterior-Posterior Dimension (3 regions of interest [ROI]: fronto-central [F7-F3-FC5 and F8-F4-FC6], temporal [T7-C3-CP5 and T8-C4-CP6], and temporo-parietal [P7-P3-PO3 and P8-P4-PO4]); and Electrodes (3) were included for lateral analysis.

Music2603_08 1/13/09 12:05 PM Page 251

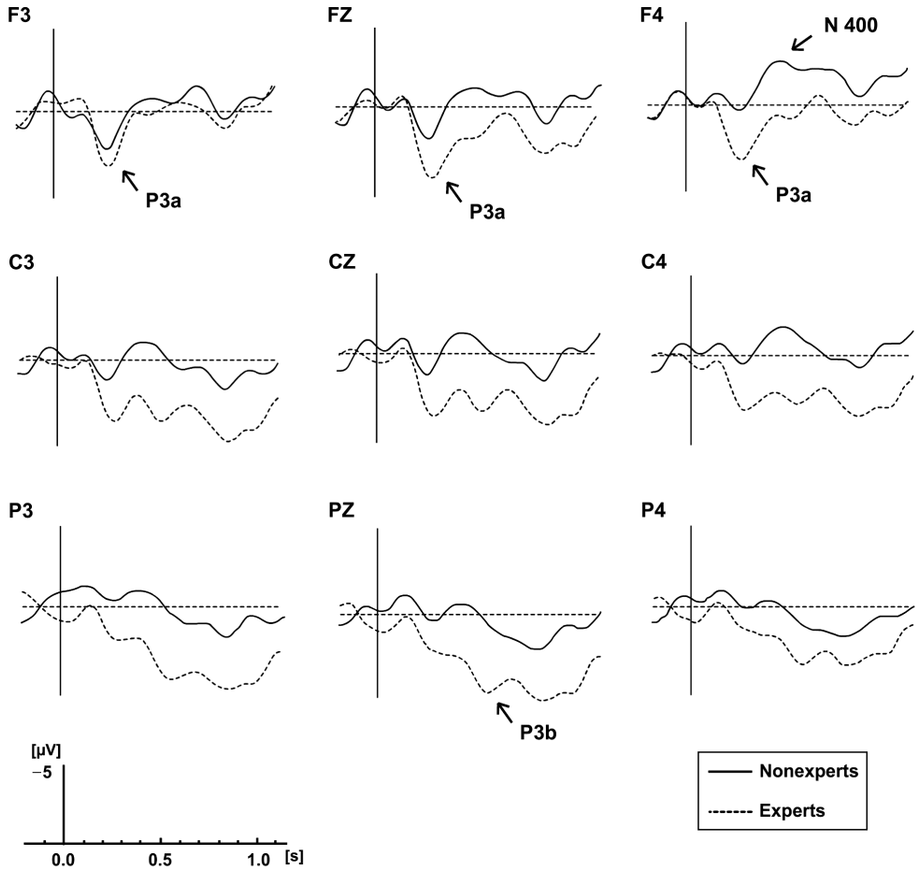


FIGURE 1. Grand average ERPs (10 participants) to incongruous targets recorded from nonexperts (solid line) and experts (dotted line). Recordings are presented for the midline electrodes (Fz, Cz, Pz) and for selected lateral electrodes representative of each ROI. Negative is up.

Results

Behavioral Data

An ANOVA including the factors Expertise and Congruity was computed on the error rate.[4] There was a main effect of Expertise, $F(1, 18) = 10.00$, $p < .005$; nonexperts made twice as many errors ($M = 30.83\%$, $SD = 4.40$) compared to experts ($M = 15.83\%$, $SD = 4.34$). Moreover, although the main effect of Congruity was not significant ($p > .64$), the Expert × Congruity interaction was significant, $F(1, 18) = 5.74$, $p < .05$. Posthoc comparisons showed that nonexperts made significantly more errors ($M = 35.00\%$, $SD = 3.07$) than experts ($M = 9.66\%$, $SD = 2.07$) for congruous targets, but no between-groups differences were found for incongruous targets (nonexperts: $M = 26.22\%$, $SD = 5.11$; experts: $M = 22.00\%$, $SD = 5.16$).

However, if the task is considered as an incongruity detection task, results show that experts (78%) and nonexperts (74%) demonstrated comparable hit rates, but that the false alarm rate was lower for experts (10%) compared to nonexperts (35%).

Electrophysiological Data

No significant effects were found both in the 0-150 ms and 150-300 ms latency bands at either midline or lateral electrodes.

In the 300-600 ms latency window, the main effect of Expertise was significant at both central and lateral electrodes, $F(1, 18) = 10.79$, $p < .005$ and $F(1, 18) = 4.83$, $p < .05$ respectively, with a larger positive effect for experts (central electrodes = 2.08 μV; lateral electrodes = 1.38 μV) than nonexperts (central electrodes = −0.54 μV; lateral electrodes = −0.19 μV); see Figure 1. Moreover, the Expertise × Hemisphere × ROI interaction also was significant, $F(2, 36) = 3.61$, $p < .05$; the

[4]Reaction times (RT) were not analyzed since participants were asked to await the end of the target TSUs to press the response button.

Music2603_08 1/13/09 12:05 PM Page 252

TABLE 1. Mean Amplitudes (μV) of Experts and Nonexperts in the 300-600 ms Latency Band for each ROI in the Left and Right Hemispheres

Expertise	Experts						Nonexperts					
HEM	LH			RH			LH			RH		
ROI	*f-c*	*t*	*t-p*	*f-c*	*t*	*t-p*	*f-c*	*t*	*t-p*	*f-c*	*t*	*t-p*
μV	0.25	1.87	2.08	0.92	1.69	1.51	0.09	0.00	–0.24	–0.51	–0.20	–0.24

Note: f-c = fronto-central, t = temporal, t-p = temporo-parietal

differences between experts and nonexperts were largest over the frontal region of the right hemisphere (see Table 1).

Estimated current source density (CSD) maps of the scalp topography of the ERP in the 300-600 ms latency range were computed for incongruous TSU targets. As can clearly be seen in Figure 2 (see color plate section), the CSD for nonexperts has a pronounced negative focus over the right frontal region. For experts, a CSD positivity develops that is distributed over the left parietal regions. Thus these topographic differences clearly show that the configuration of brain sources active during listening to incongruous TSUs is different for nonexperts and experts. In contrast, no differences linked to expertise were found for congruous TSU targets (Figure 3).

Finally, while the main effect of Expertise was not significant, $F(1, 18) = 3.26$, $p = .08$, in the 600-1000 ms latency band, there was somewhat larger positivity for experts (2.39 μV) compared to nonexperts (0.41 μV).

Discussion

The primary motivation behind this experiment was to determine whether the musical excerpts defined as TSUs indeed have specific conceptual content. To this end, we created incongruous TSUs to determine whether the shifting in midstream of a TSU category (e.g., fall) into a different TSU category (e.g., stationary) would be noticed both by TSU experts and nonexperts and, in such case, whether they would elicit N400- or P300-type components.

In line with our hypothesis, behavioral data revealed that the level of performance of TSU experts to decide whether the TSUs were congruous or incongruous was significantly higher than for nonexperts. This is to be expected insofar as experts have theoretical knowledge about the different TSUs categories they can apply to the new TSUs presented here. The finding of a significant

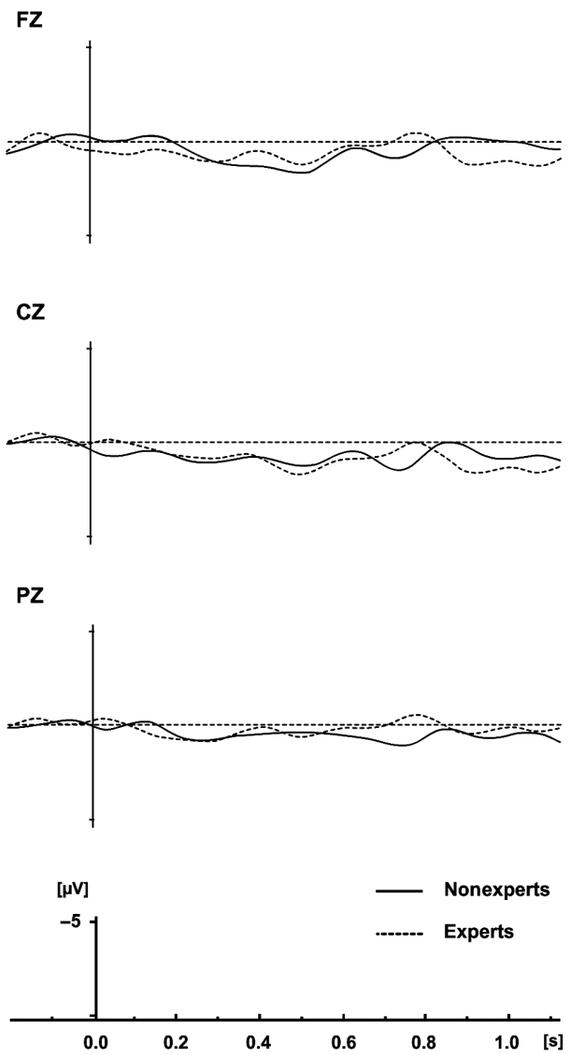


FIGURE 3. Grand average ERPs (10 subjects) to congruous targets, recorded from nonexperts (solid line) and experts (dotted line). Recordings are presented for midline electrodes (Fz, Cz, Pz). Effects were similar (and not significant) at lateral electrodes.

Music2603_08 1/13/09 12:05 PM Page 253

Expertise by Congruity interaction is also in line with this interpretation. It shows that, when the TSUs matched with previous knowledge acquired in electroacoustic classes (as is the case for congruous TSUs), experts performed better than nonexperts.

Moreover, while both experts and nonexperts detected incongruous TSUs equally well—as reflected in similar hit rates—nonexperts were more likely to consider congruous TSUs as incongruous, as reflected by their higher false alarm rate. Clearly, the nonexperts used a less stringent criterion for calling something incongruous than experts. Again, this is to be expected given their unfamiliarity with the congruous stimuli. What remains to be determined, however, is whether incongruity detection occurs on the basis of perceptual or conceptual changes. ERP data may shed some light on this issue.

Whereas the brain activity elicited by congruous TSUs was very similar for experts and nonexperts—suggesting that similar processes were at play for both groups—qualitative differences emerged in the ERPs to incongruous TSUs. Around 300 ms after an incongruous TSU started to transform into another one, positive-going ERPs were found in experts. By contrast, a negative component developed with maximum amplitude at 400 ms in nonexperts. Detailed scalp topography analyses revealed that the ERP differences between experts and nonexperts were largest over the right frontal regions. Moreover, computation of CSD maps (see Figure 2) showed a positive focus over the left parietal region for experts and a negative focus over the right frontal region for nonexperts. These topographic differences clearly showed that the configuration of brain sources active during listening to incongruous TSUs differed for nonexperts and experts. Consequently, incongruous TSUs most likely elicited different processes in both groups.

Considering first the congruous TSUs, behavioral data showed that the level of performance for congruous TSUs was higher for experts than nonexperts. However, no differences were found in the ERPs (that did not differ from baseline; see Figure 3), probably because all TSUs were different from each other and each one referred to a different concept. Moreover, since no specific events occurred in the congruous TSUs, no specific variations were elicited and synchronized to be visible in the ERPs.

By contrast with congruous TSUs, incongruous TSUs did elicit different ERPs in experts and nonexperts. Based upon previous ERP literature, one may argue that the long lasting positivity found in experts encompasses different ERP components (Hansenne, 2000). The first positive peak, around 300 ms and of maximum amplitude over the frontal regions (see Figure 1), is likely to be a "novelty" P300, also called the P3a component (Courchesne et al., 1975). Since its first description, the P3a has been related to orienting and to the momentary shift of attention toward an unexpected perturbation in the environment (Yamaguchi & Knight, 1991). Interestingly, additional ERP and MRI findings also indicate frontal lobe activity for the detection of rare but alerting stimuli (McCarthy, Luby, Gore, & Goldman-Rakic, 1997). Thus, the presence of a P3a component is likely to reflect the fact that incongruous transformations of the TSUs very rapidly and possibly automatically attracted the attention of experts. They were surprised by the change. Following the P3a, a second positive peak, around 550 ms, and of maximum amplitude over parietal regions (see Figure 1) is likely to be a P3b (Comerchero & Polich, 1999; Donchin, 1981; Kutas, Neville, & Holcomb, 1987; Polich & Criado, 2006). The P3b is thought to be evoked whenever events force an updating of stimulus representations held in working memory (Donchin & Coles, 1988), may be through top-down controlled attentional processes (Debener, Kranczioch, Herrmann, & Engel, 2002). Thus, the occurrence of a P3b is likely to reflect the conscious categorization of the TSU as incongruous. By contrast, whereas the P3a was also elicited in nonexperts (see Figure 1, Fz and F3), this first positive component was followed by a negative component of maximum amplitude over the right frontal region around 400 ms (see Figure 1, F4, and Figure 2). Thus, similar levels of performance for experts (78% correct) and nonexperts (74%) can be achieved using different processes, as reflected by the ERPs.

An important point that remains to be discussed is whether participants did consider that TSUs were incongruous on the basis of perceptual or conceptual changes. The finding that P3a components are elicited both in experts and nonexperts shows that the incongruous TSU transformation was perceptually noticeable and automatically attracted attention in both groups. Thus, perceptual differences linked to the greater heterogeneity of incongruous compared to congruous TSUs certainly influenced the decision. However, following the P3a, different processes seemed to be called into play in both groups. For experts, the first automatic attentional shift induced by the perceptually incongruous transformation of the TSU (and reflected by the P3a) was followed by a P3b component that most likely reflected the decision that the TSU was heterogeneous and included a change that was unexpected based on their specific knowledge of TSU categories. In contrast, in nonexperts the P3a component

Music2603_08 1/13/09 12:05 PM Page 254

was followed by a negative component (note that this negative component probably overlapped the P3a, thereby explaining why P3a amplitude was smaller for nonexperts than experts).

While these differences between experts and nonexperts may seem counterintuitive, they may reflect differences in task difficulty. Because the task was easy for experts, they categorized the TSUs rapidly, as reflected by the occurrence of P3b component that may overlap and mask the negativity that developed in nonexperts. From a theoretical point of view, the main question is whether the negative component found in nonexperts can be considered an N400 component.

In the nonexpert group the negative component developed in the typical latency band of the N400 component (between 300 and 600 ms) over right-frontal regions. Because the N400 component has been found to be elicited by written words (e.g., Kutas & Hillyard, 1980), spoken words (e.g., McCallum, Farmer, & Pocock, 1984) and American sign language (e.g., Neville et al., 1997), it has been argued that its occurrence is relatively independent of sensory input modality (Kutas et al., 1987). The results of several studies nevertheless showed differences in N400 scalp distribution, suggesting that the N400 may not be completely modality independent. For instance, the N400 elicited by unrelated pictures in word-picture priming experiments shows a more anterior distribution than for words (Ganis, Kutas, & Sereno, 1996), and is sometimes larger over the left than the right hemisphere (Holcomb & McPherson, 1994). More directly related to our stimuli, the N400 elicited by unrelated environmental sounds in word-sound priming experiments typically is larger over the frontal regions of the left hemisphere (whereas N400s to words in sound-word priming designs are larger over the centro-parietal regions of the right hemisphere; Plante et al., 2000; Van Petten & Rheinfelder, 1995). These differences in scalp distribution therefore suggest that there may be a family of N400 components, each one sensitive to different modalities of input but with similar functional significance.

As mentioned in the introduction, the N400 component typically is considered as reflecting semantic processing. In this framework, the negative component elicited in nonexperts may reflect the semantic-conceptual processing of incongruous TSUs. In a seminal experiment, Koelsch et al. (2004) have reported that musical excerpts can prime subsequently presented words in nonmusicians so that words unrelated to the semantic or conceptual meaning of the musical excerpts are associated with N400 components. Moreover, Steinbeis and Koelsh (2008) have shown that violations of tension-resolution patterns in tonal music elicit N400-like components in nonmusicians (N5). Thus, our results with nonexperts may be taken to complement and extend these findings by showing that N400s also can be elicited by incongruous changes in the music itself, without any words involved. Finally, Picton (1988) suggested that the amplitude of the N400 varies with the amount of memory search necessary to find the meaning of a stimulus. In this framework, larger N400s are expected in nonexperts than in experts because the memory search for concepts is certainly more difficult for the former than for the later.

In conclusion, the present results show that incongruous TSUs are most likely considered as such on the basis of both perceptual processes (as reflected by the occurrence of a P3a in both groups) and conceptual processes, if the negative component elicited in nonexperts can be considered as an N400-like component. The right-frontal distribution of the present effect may suggest hemispheric differences in the neural networks underlying the processing of words and TSUs. However, right anterior negativities (RAN) have also been described in previous experiments using chord sequences when the final chords does not fit within the tonal context (Koelsch, Gunter, Friederici, & Schröger, 2000; Patel, Gibson, Ratner, Besson, & Holcomb, 1998). The changes from one TSU category to another occurring in incongruous TSUs cannot be considered as tonal changes, Nevertheless, it may be that RAN also may be elicited by the processing of perceptually heterogeneous stimuli. Clearly, these scalp distribution differences need to be better understood, perhaps by using other brain imaging methods that render better spatial resolution than ERPs. Magnetoencephalography or fMRI may locate more precisely the processing of semantic aspects of music. This study should be considered as a first step in using music-music priming designs to better understand the similarities and differences between linguistic and musical semantics.

Author Note

We thank the members of the Laboratoire Musique et Informatique de Marseille for fruitful discussions and for their help in testing the stimuli, as well as all the participants in the study. This work was made possible thanks to a grant from the Ministère de l'éducation nationale et de la recherche, ACI "Terrain, Techniques et Théories" (045692) to Mireille Besson and from the Ministère de l'éducation nationale et de la recherche (Fonds National

Music2603_08 1/13/09 12:05 PM Page 255

pour la Science—TTT4014) to Prof Charles Tijus. Aline Frey is supported by the Laboratoire Cognitions Humaine et Artificielle and Céline Marie is supported by a fellowship from the Ministère de l'éducation nationale et de la recherche.

Correspondence concerning this article should be addressed to Aline Frey, Laboratoire CHART: Cognitions Humaine et Artificielle—EA 4004—UFR 7 2, rue de la Liberté, 93 526 Saint-Denis, Cedex, France. E-MAIL: aline.frey@cognition-usages.org

References

BARRETT, S. E., & RUGG, M. D. (1989) Event-related potentials and the semantic matching of faces. *Neuropsychologia, 27*, 913-922.

BARRETT, S. E., & RUGG, M. D. (1990). Event-related potentials and the semantic matching of pictures. *Brain and Cognition, 14*, 201-212.

BICKNELL, J. (2002). Can music convey semantic content? A Kantian approach. *The Journal of Aesthetics and Art Criticism, 60,* 253-261.

BIGAND, E., & POULIN-CHARRONNAT, B. (2006). Are we "experienced listeners"? A review of the musical capacitites that do not depend on formal musical training. *Cognition, 100,* 100-130.

BOOTZ, P., & HAUTBOIS, X. (2007a). Les motifs temporels paramétrés [Configuration of temporal motives]. 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"* [Toward a general semiotics of time in arts. Proceeding of "Temporal Semiotic Units (TSU), new tools for musical analyses: Theory and applications"] (pp. 139-176), Sampzon: Delatour-France.

BOOTZ, P., & HAUTBOIS, X. (2007b). Time 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.

CHABY, L., JEMEL, B., GEORGE, N., RENAULT, B., & FIORI, N. (2001). An ERP study of famous face incongruity detection in middle age. *Brain and Cognition, 45*, 357-377.

COMERCHERO, M. D., & POLICH, J. (1999). P3a and P3b from typical auditory and visual stimuli. *Clinical Neurophysiology, 110*, 24-30.

COURCHESNE, E., HILLYARD, S. A., & GALAMBOS, R. (1975). Stimulus novelty, task relevance and the visual evoked potential in man. *Electroencephalography and Clinical Neurophysiology, 39*, 131-143.

CUMMINGS, A., CEPONIENE, R., KOYAMA, A., SAYGIN, A. P, TOWNSEND, J., & DICK, F. (2006). Auditory semantic networks for words and natural sounds*. Brain Research, 1115*, 92-107.

DANIÉLOU, A. (1993). *Sémantique musicale. Essai de psychophysiologie auditive* [Musical semantics: An essay on auditory psychophysiology]. Paris: Hermann.

DEBENER, S., KRANCZIOCH, C., HERRMANN, C. S., & ENGEL, A. K. (2002). Auditory novelty oddball allows reliable distincion of top-down and bottum-up processes of attention. *International Journal of Psychophysiology, 46*, 77-84.

DELALANDE, F., FORMOSA, M., FRÉMIOT, M., GOBIN, P., MALBOSC, P., MANDELBROJT, J., & PEDLER, E. (1996). *Les unités sémiotiques temporelles—Éléments nouveaux d'analyse musicale* [Temporal semiotic units—New elements for musical analysis; book-CD]. Marseille: MIM/Documents Musurgia.

DONCHIN E. (1981). Surprise . . . surprise? *Psychophysiology 18*, 493-513.

DONCHIN, E., & COLES, M.G. H. (1988). Is the P300 component a manifestation of cognitive updating? *Behavioral and Brain Sciences, 11*, 357-427.

DOWLING, W. J., & HARWOOD, D. (1986). *Music cognition.* New York: Academic Press.

FRANCÈS, R. (1958). *La perception de la musique* [Music perception]. Paris: Vrin.

GANIS, G., KUTAS, M., SERENO, M. I. (1996). The Search for "Common Sense": An Electrophysiological Study of the Comprehension of Words and Pictures in Reading. *Journal of Cognitive Neuroscience, 8(2),* 89-106.

GASER, C., & SCHLAUG, G. (2003). Brain structures differ between musicians and non-musicians. *Journal of Neuroscience, 23*, 9240-9245.

HANSENNE, M. (2000). Le potentiel évoqué cognitif P300 (I): Aspects théorique et psychobiologique [Cognitive evoked potential P300 (I): Theoretical and psychobiological aspects]. *Neurophysiologie Clinique, 30*, 191-210.

HOLCOMB, P. J., & MCPHERSON, W. B. (1994). Event-related brain potentials reflect semantic priming in an object-decision task. *Brain and Cognition, 24*, 259-276.

JONES, M. R. (1981). Music as a stimulus for psychological motion: Part 1. Some determinants of expectancies. *Psychomusicology, 1*, 34-51.

JONES, M. R. (1982). Music as a stimulus for psychological motion: Part 2. An expectancy model. *Psychomusicology, 2*, 1-13.

KOELSCH, S., SCHRÖGER, E., & TERVANIEMI, M. (1999). Superior attentive and pre-attentive auditory processing in musicians. *NeuroReport, 10*, 1309-1313.

KOELSCH, S., GUNTER, T. C, FRIEDERICI, A. D., & SCHRÖGER, E. (2000). Brain indices of music processing: "Nonmusicians" are musical. *Journal of Cognitive Neuroscience, 12,* 520-541.

KOELSCH, S., KASPER, E., SAMMLER, D., SCHULZE, K., GUNTER, T., & FRIEDERICI, A. D. (2004). Music, language

Music2603_08 1/13/09 12:05 PM Page 256

and meaning: Brain signatures of semantic processing. *Nature Neuroscience, 7,* 302-307.

KOYAMA S., NAGEISHI Y., & SHIMOK-CHI M. (1992). Effects of semantic context and event-related potentials: N400 correlates with inhibition effect. *Brain and Language, 43*, 668-681.

KUTAS, M., & HILLYARD, S. A (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 207*, 203-205.

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

KUTAS, M., NEVILLE, H. J., & HOLCOMB, P. J. (1987). A preliminary comparison of the N400 response to semantic anomalies during reading, listening, and singing. *Electroencephalography and Clinical Neurophysiology, 39,* 325-330.

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

LERDAHL, F., & JACKENDOFF, R. (1983). *A generative theory of tonal music*. Cambridge, MA: MIT Press.

MCADAMS, S. (1989), Les nombreux visages de la cognition humaine dans la recherche et la pratique musicales [The numerous aspects of human cognition in musical research and practice]. In P. Mardage (Ed.), *La musique et les sciences cognitive* [Music and cognitive sciences] (pp. 11-20). Bruxelles: Editions Pierre Mardaga.

MCCALLUM, W. C., FARMER, S. F., & POCOCK, P. V. (1984). The effects of physical and semantic incongruities on auditory event-related potentials. *Electroencephalography and Clinical Neurophysiology*, *59*, 477-488.

MCCARTHY, G., LUBY, M., GORE, J., & GOLDMAN-RAKIC, P. (1997). Infrequent events transiently activate human prefrontal and parietal cortex as measured by functional MRI. *Journal of Neurophysiology, 77,* 1630-1634.

MEYER, L. B. (1956). *Emotion and meaning in music*. Chicago: University of Chicago Press.

NEVILLE, H. J., COFFEY, S. A., LAWSON, D. S., FISCHER, A., EMMOREY, K., & BELLUGI, U. (1997). Neural systems mediating American sign language: Effects of sensory experience and age of acquisition. *Brain and Language, 57*, 285-308.

ORGS, G., LANGE, K., DOMBROWSKI, J-H., & HEIL, M. (2006). Conceptual priming for environmental sounds and words: An ERP study. *Brain and Cognition, 62,* 267-272.

ORGS, G., LANGE, K., DOMBROWSKI, J-H., & HEIL, M. (2007). Is conceptual priming for environmental sounds obligatory? *International Journal of Psychophysiology, 65,* 162-166.

PATEL, A. D., GIBSON, E., RATNER, J., BESSON, M., & HOLCOMB, P. J. (1998). Processing syntactic relations in language and music: An event-related potential study. *Journal of Cognitive Neuroscience, 10,* 717-733.

PICTON, T. W. (1988). The endogenous evoked potentials. In E. Basar (Ed.), *Dynamics of sensory and cognitive processes in the brain* (pp. 258-272). New York: Springer-Verlag.

PLANTE, E., VAN PETTEN, C., & SENKFOR, A. J. (2000). Electrophysiological dissociation between verbal and nonverbal semantic processing in learning disabled adults. *Neuropsychologia, 38*, 1669-1684.

POLICH, J., & CRIADO, J. R. (2006). Neuropsychology and neuropharmacology of P3a and P3b. *International Journal of Psychophysiology, 60,* 172-185.

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*, 27-52.

SCHAEFFER, P. (1966), *Traité des objets musicaux* [Treatise on musical objects]. Paris: Editions du Seuil.

SEUNG, Y., KYONG, J. S., WOO, S. H., LEE, B. T., & LEE, K. M. (2005). Brain activation during music listening in individuals with or without prior music training. *Neuroscience Research, 52*, 323-329.

STEINBEIS, N., & KOELSCH, S. (2008). Shared neural resources between music and language indicate semantic processing of musical tension-resolution patterns. *Cerebral Cortex*, *18*, 1169-1178.

SUTTON, S., BRAREN, M., ZUBIN, J., & JOHN, E. R (1965). Evoked potential correlates of stimulus uncertainty. *Science, 150*, 1187-1188.

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.

VAN PETTEN, C., & RHEINFELDER, H. (1995). Conceptual relationships between spoken words and environmental sounds: Event-related brain potential measures. *Neuropsychologia, 33*, 485-508.

YAMAGUCHI, S., & KNIGHT, R. T. (1991). Anterior and posterior association cortex contributions to the somatosensory P300. *Journal of Neuroscience, 11,* 2039-2054.