



HAL
open science

Influence of regular rhythmic versus textural sound sequences on semantic and conceptual processing

Laure-Hélène Canette, Philippe Lalitte, Barbara Tillmann, Emmanuel Bigand

► **To cite this version:**

Laure-Hélène Canette, Philippe Lalitte, Barbara Tillmann, Emmanuel Bigand. Influence of regular rhythmic versus textural sound sequences on semantic and conceptual processing. *Music Perception*, In press. hal-03384384

HAL Id: hal-03384384

<https://hal.science/hal-03384384>

Submitted on 18 Oct 2021

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.

Influence of regular rhythmic versus textural sound sequences
on semantic and conceptual processing

Laure-Hélène Canette^{1,2}, Philippe Lalitte^{1,3}, Barbara Tillmann² * & Emmanuel Bigand^{1*}

¹The Laboratory for Research on Learning and Development, CNRS-UMR 5022, University of Burgundy, Dijon, France.

²Lyon Neuroscience Research Center, Auditory Cognition and Psychoacoustics Team, CNRS-UMR 5292, Inserm U 1028, University of Lyon 1, Bron, France.

³Research Institute in Musicology (IReMus), CNRS-UMR 8223, Sorbonne University, Paris, France.

*Shared last authors

ABSTRACT

Conceptual priming studies have shown that listening to musical primes triggers semantic activation. The present study further investigated with a free semantic evocation task (1) how rhythmic vs. textural structures affect the amount of words evoked after a musical sequence and (2) whether both features also affect the content of the semantic activation. Rhythmic sequences were composed of various percussion sounds with a strong underlying beat and metrical structure. Textural sound sequences consisted in blended timbres and sound sources evolving over time without identifiable pulse. Participants were asked to verbalize the concepts evoked by the musical sequences. We measured the number of words and lemmas produced after having listened to musical sequences of each condition, and we analysed whether specific concepts were associated with each sequence type. Results showed that more words and lemmas were produced for textural sound sequences than for rhythmic sequences and that some concepts were specifically associated with each musical condition. Our findings suggest that listening to musical excerpts emphasizing different features influences semantic activation in different ways and extent. This might possibly be instantiated via cognitive mechanisms triggered by the acoustic characteristics of the excerpts as well as the perceived emotions.

KEY WORDS

Semantics, Music, Rhythm, Texture, Conceptual priming.

INTRODUCTION

Listening to music can trigger listeners to think about specific concepts, ideas, events or memories. This phenomenon relates to “extra-musical meaning”, in opposition to “intra-musical meaning” (i.e., structural reference of one musical element to at least one other musical element) (e.g. Koelsch, 2011; Meyer, 1956; Patel, 2008). Following Peirce (1931-58), extra-musical meaning has been described to include three dimensions (e.g., Koelsch, 2011): (1) iconic musical meaning (i.e., meaning emerging from music or musical features reminding the listeners of sounds of objects or qualities of objects and abstract concepts); (2) indexical musical meaning (i.e., meaning from music expressing a psychological state of an individual, for example an emotion); (3) symbolic musical meaning (i.e., meaning that arises from explicit or conventional extra-musical association, for example, the national anthem). Schlenker (2019) also supports the idea that music conveys information about a music-external reality. He argues that music semantics aggregates two types of inferences: (1) inferences from non-musical auditory cognition (i.e., inferences about the “virtual sources” of the music), and (2) inferences from properties of the tonal pitch space (i.e., how voices evolve in tonal pitch space). The first kind of inferences could pertain to the type of sources, their potential actions and their location, notably on the basis of timbre, speed and speed modifications, loudness and pitch height for example. The second type of inferences refer to music-specific elements (e.g., musical tension, dissonance, cadences), it depends on the musical idiom considered, and thus varies across cultural traditions.

Extra-musical meaning conveyed by music has been investigated with a priming paradigm: a musical sequence (the prime) influences semantic processing of a subsequently presented target word. Koelsch et al. (2004) reported a larger N400 (i.e., an electrophysiological marker for semantic integration) for target words when they were semantically incongruous with the presumed meaning or semantics of the previously

presented musical prime rather than semantically congruous. In a behavioural follow-up study, Poulin-Charronnat, Grieser, Meyer and Koelsch (2006) showed that reaction times in a lexical decision task were also influenced by the semantic relatedness between musical primes and the subsequently presented target words. The semantic priming effect has also been reported with short musical excerpts: Daltrozzo and Schön (2009a, 2009b) reported a N400 priming effect for target words following musical excerpts of 1 second. With this short duration, a musical excerpt contains mostly spectral information (often referred to as “musical timbre”). According to the authors, musical timbre as well as the energy, tension and arousal carried by a sound or series of sounds might be rich and informative enough to trigger semantic activation (Daltrozzo & Schön, 2009a). Grieser Painter and Koelsch (2011) provided further evidence about the potential influence of timbre, notably by reporting a N400 priming effect triggered by various synthesized timbres, resembling or not musical instruments, outside of a musical context. They concluded that a musical context is not necessary for basic musical features, such as timbre, to activate meaningful mental, conceptual representations. They also showed that time of exposure and depth of processing are important factors influencing the elicitation of meaningful associations between a musical sound prime and subsequently presented target words. Their study confirmed the hypothesis that low-level, acoustic features of music are extracted in early processing stages and are directly linked to meaningful conceptual representations (as also proposed by a model of music perception, Koelsch & Siebel, 2005). This link can be also found in the framework proposed by Schön, Ystad, Kronland-Martinet and Besson (2009) to explain conceptual priming effects between music (or sounds in general) and language. This framework postulated an “amodal concept lexicon” that could be activated by early processing steps of acoustic feature processing, and connects music and language processing. It is an interesting framework for cross-domain priming studies, and we will discuss this framework further in link with our data in the

discussion section (even though our experiment did not directly aim to test this framework).

Focussing on more emotional aspects, Sollberger, Reber and Eckstein (2003) manipulated the consonance/dissonance of prime chords and their affective congruence with the subsequently presented target word. Using a word-evaluation task (i.e., to judge whether the target word refers to a positive or negative concept), this study reported longer response times after an affectively incongruent chord prime (e.g., a consonant prime chord followed by an affectively negative target word) than after an affectively congruent chord prime (e.g. a consonant prime chord followed by an affectively positive target word). Congruent results were reported by Steinbeis and Koelsch (2011) who manipulated three musical parameters: consonance/dissonance, mode (major/minor) and timbre. For each of these parameters, longer response times and the N400 were observed for affectively incongruent pairs compared to affectively congruent pairs.

The potential relation between the musical features of the prime and the semantic content of the target has been also investigated for iconic musical meaning in particular. Using pictures as targets, Zhou, Jiang, Wu and Yang (2015) reported that the concepts of movement (i.e., in motion or at rest) and space (i.e., open or closed space) rely on different musical features: temporal features (i.e., tempo and spectral flux) predicted the expression of movement, whereas pitch range and intensity flux predicted the expression of space. Musical features can also influence cognitive processing, such as mental representation and categorization level, for example (Hansen & Melzner, 2014): The sound stimuli that included cues to distance (reverberation and novel harmonic combinations) and abstraction (low segmentation) triggered more abstract construal, whereas sound stimuli that included more acoustic cues to proximity (no reverberation and familiar harmonic combinations) and concreteness (high segmentation) triggered more concrete construal. The authors highlighted that future studies should further explore more musical cues, notably timbre.

Other recent evidence of the differential influence of musical features on concept activation and cognitive processing has been reported for children for speech perception and production (Canette et al., 2020a). This study compared the effects of rhythm and timbre (i.e., implemented in two types of musical sequences) on two linguistic tasks. The regular rhythmic sequences had a strong underlying beat and metrical structure, while textural sound sequences exhibited spectro-temporal changes and progressions, but were without underlying pulse. The results showed that listening to rhythmic musical sequences enhance subsequent syntax processing (i.e., grammatical judgments on auditorily presented sentences) compared to textural sound sequences, whereas textural sound sequences promote verbalization about the music more strongly than do rhythmic sequences. This result may be understood as reflecting greater semantic network activation for textural sound sequences.

Overall, previous studies on extra-musical meaning have suggested that listening to music activates semantic concepts, and that these concepts would be associated to the musical characteristics. Our present study aimed to further investigate whether semantic activation depends on the musical parameters, in terms of how extended his semantic activation could be and also in in terms of content. For this aim, we investigated the effect of textural sound sequences vs. rhythmic sequences on lexical activation, with a free semantic evocation task. In this task, participants were asked to verbalize concepts evoked by a given musical sequence heard previously. We made the hypothesis that the number of words and lemmas evoked, as well as the semantic content of the verbal productions, could be modulated by the characteristics of the heard music. Unlike our previous study (Canette et al., 2020a), we here focused on an adult population, aiming to confirm our hypotheses in a mature lexical and linguistic system, and in a population for whom the task should be easier. Based on previous studies showing the effectiveness of timbre to convey concepts (Daltrozzo & Schön, 2009a, 2009b; Grieser Painter & Koelsch, 2011; Koelsch et al., 2004) and on the results obtained

with children (Canette et al., 2020a), we hypothesized that (1) textural sound sequences should promote verbalization (i.e., more words and lemmas produced) more strongly than the regular rhythmic sequences, interpreted as a greater semantic network activation¹, and that (2) the semantic content of the productions (related to iconic and indexical musical meaning) should be influenced by the previously heard musical characteristics of the sequence type (texture vs. rhythm) such as rhythmic characteristics, timbre, emotional aspects. In addition, as a methodological control, the participants judged the musical material for the level of pleasure, interest and emotion using subjective scales, allowing us to further investigate the obtained production performance.

MATERIAL AND METHODS

1. Participants. Twenty-five native French-speaking young adults participated in the study² (mean = 21.28 years, SD = 2.89, range = 19 – 32; 19 women and 6 men). Average years of formal musical instrument training was 2.13 (SD = 3.52) for the entire group: Ten participants had musical instrument training on average for (5.67 years, SD = 3.61), while the others reported 0.

¹ We need to acknowledge that the words produced by the participants could be words directly evoked by the prime as well as words used by the participants to post-hoc describe or react to the musical sequence. Even though with our current methodology, we cannot distinguish between the two, we make the hypothesis that also reactions or post-hoc descriptions should be related to the evocations of the prime.

² The group size corresponds to group sizes used in numerous research in cognitive psychology, including conceptual priming studies: for example, groups between 16 and 24 in Koelsch et al. (2004) and groups of about 20 participants in Daltrozzo & Schön, 2009a and 2009b; Grieser Painter & Koelsch, 2011; Steinbeis & Koelsch, 2011; Zhou et al., 2015; Schön et al., 2009). Moreover, the present study builds on Canette et al., 2020a testing 14 typically developing children on the same semantic evocation task: the effect size for the number of content words was $d_z = 0.77$, and the power was 0.76 (two-tailed test, with $\alpha = 0.05$; calculated with G*Power). In the present study on adults, we expected about the same effect size, and we increased the number of participants to reach higher power. Note that the results confirmed this, as we observed similar effect size ($d_z = 0.82$) and increased power (0.98) (two-tailed test, with $\alpha = 0.05$; calculated with G*Power) for the dependent variable *number of content words*.

The study was performed in an educational context, in which students, in exchange for course credits, participate in a non-invasive laboratory experiment. They took part naïvely in the study, and after participation, they could ask the experimenter about the purpose of the study to further their understanding of experimental psychology. Informed written consent was obtained from all participants prior to taking part in the experiment. The study was anonymous and fully obeyed to the Helsinki Declaration, Convention of the Council of Europe on Human Rights and Biomedicine.

2. Material

Musical stimuli. Two types of musical sequences of 17 seconds were used: regular rhythmic sequences and textural sound sequences (see Appendix 1 for more details, and supplemental material 1 for auditory examples).

Six regular rhythmic sequences were constructed on the basis of a 4/4 meter and a 120 BPM tempo, that is with an underlying inter-beat-interval of 500 ms (i.e., 2 Hz). These sequences were formed by 4 layers of rhythmic patterns, leading to a four-level rhythmic hierarchy. To create variety across the sequences beyond the temporal pattern, each layer was played by one (and sometime two) percussion instrument(s) (i.e., bass drum, snare drum, tom-tom, or cymbal). The rhythmic sequences were composed by one of the authors without automatic generation algorithm by playing rhythms on a MIDI Keyboard linked to a computer. They were composed in an electro/dance music genre.

Textural sound sequences were extracted from contemporary art music. They were characterized by the blending of continuous melodic or harmonic strata evolving across time, without regular rhythmic patterns or underlying pulse or meter. Texture in music can be defined as a “grain, web, surface or nap” (Smith-Brindle, 1966), which is more or less homogenous resulting from the merging of separate elements. Rowell (1983) used the

metaphor of weaving to conceive texture as an aesthetic surface where the warp threads represent pitch, vertical axis, and simultaneity, while the woof represents time, the horizontal axis, and sequence. Texture is working as a gradual transformation, it forms a sound matter produced by the fusion of many instruments into a slowly evolving texture. In our study, we used smooth textures with no identifiable pulse or meter. In these textures, parameters like intervals and rhythm are less relevant than sound fields determined by the interaction of different parameters, such as density, timbre or dynamic (Lalitte, 2003).

All musical sequences were taken from real recordings, and normalized in loudness with the software Audacity (level of normalization applied in the plug-in: -5 dB).

3. Procedure. Participants were asked to listen to a musical sequence and then to tell what the music made them think about and what they imagined during listening to the music (i.e., words, images, concepts)³. There was no time limit, and the experimenter launched the next musical sequence when the participant has finished with the production. Participants were free to indicate if nothing came in their mind for a given musical sequence (e.g., skipping the response or saying “I don’t know”). The order of the musical sequences was pseudo-randomly determined for each participant with the constraint to avoid listening to the same sequence type more than twice consecutively. Half of the participants started with a regular rhythmic sequence and half with a textural sound sequence. Participants worked on all 12 sequences (i.e., 6 regular rhythmic sequences and 6 textural sound sequences).

After the semantic evocation task, the musical sequences were presented again (in the same pseudo-random order), and participants were asked to judge each sequence on three different subjective scales from 1 to 5. The first scale assessed perceived emotion (from “very

³ « For each music, listen carefully, and at the end, describe orally what the music made you think about, what you imagined during listening to the music (what concepts, words, pictures come to mind as you listen to the music) ».

sad” to “very happy”), the second scale assessed interest (from “very boring” to “very interesting”), and the third scale assessed pleasure (from “I don’t like at all” to “I like very much”). Participant could respond during listening to the music by selecting one of the five possibilities displayed on the screen (the response stopped the presentation of the soundfile). We designed these judgment scales to measure whether and to what extent the participants distinguished the two music conditions (i.e., regular rhythmic vs. textural sound materials) on the three aspects, notably emotion, interest and pleasure.

The semantic evocation task and the judgment task were programmed with OpenSesame software (Mathôt, Schreij & Theeuwes, 2012). Stimuli were presented via loudspeakers (M-Audio AV40), and verbal productions were recorded using a voice recorder (Olympus).

4. Data Analysis. Participants’ productions were first transcribed. Sentences or clauses that indicated the end of the response (e.g., “that’s all”) were removed. Transcriptions were encoded in xml (extensible markup language). The resulting corpus was analyzed with the software TXM, version 0.7.8 (Heiden, 2010; for further details see <http://textometrie.ens-lyon.fr/html/doc/manual/0.7.9/fr/manual1.xhtml>), an open-source textual analysis software for textometry, and the software RStudio (RStudio Team, 2016). Part-of-speech tagging was realized in TXM with TreeTagger (Schmid, 1994), which is commonly used in most of textometric tools.

For a first analysis, the number of words (i.e., total number of occurrences or productions) and the number of lemmas (i.e., lexical unit grouping all the graphical forms, for example the two forms “child” and “children” group into the lemma “child”) were extracted in TXM for each musical sequence for each participant. This analysis focused on grammatical types of content words (nouns, verbs, adverbs, and adjectives) to investigate the lexical

structure of the corpus. Performance after the two sequence types was compared with (a) two-sided paired *t*-tests with participants (t_1) as random variable and b) two-sided independent group *t*-tests with items (t_2) as random variable. We reported effect sizes in terms of d_z or d using G*Power software (Faul, Erdfelder, Lang, & Buchner, 2007), with the following proposed categories for interpretation: small = 0.20, medium = 0.50, large = 0.80 (Cohen, 1988). To complement the *t*-test analysis, we also ran a generalized linear mixed model on the number of content words and on the number of content lemmas, respectively, aiming to investigate the results on a trial-by-trial basis while controlling for random effects. We considered musical sequence (regular rhythmic, textural sound) as the fixed effect, and participants ($N = 25$) and musical excerpt ($N = 12$) as random effects. Statistical analysis was performed with Rstudio (RStudio Team, 2016), using the package “glmmTMB” (Brooks et al., 2017). A negative binomial distribution rather than a Poisson distribution was used because the data showed over-dispersion (Ismail & Jemain, 2007). The syntax for the model was as follows:

```
Model <- glmmTMB (dependant variable ~ RhythmTexture + (1 |participants) + (1
  |musical excerpts), family="nbinom2", data=data).
```

For significance testing of fixed effects, the *Anova* function (using type III Wald chi square tests) from the *car* package (Fox & Weisberg, 2011) was used.

To analyse the semantic content of the verbal productions, two statistical analyses were run: a factorial correspondence analysis, and an analysis of specificities.

(1) The factorial correspondence analysis allows for obtaining a global picture of the data. It investigates the lexical structure of a textual corpus on the basis of a frequency table (i.e. number of occurrences for each unit in each set). The 12 musical sequences formed the 12 sets of the analysis (also referred to as « parts »). As units, we analyzed the lemmas of content words (with a minimum frequency of 2 across the corpus, leading to 399 lemmas).

The resulting graphical representation displays the distances between the sets (here: musical sequences) and the distances between the units (here: lemmas) considering their frequency across the units or across the sets, respectively. Two *sets* will be represented with small distance (i.e., close in the space) when they have similar lexical profiles, and two *units* will be represented with small distance when they have comparable textual distribution profiles. The origin of the axes of the resulting graphical representation corresponds to the average profiles, so the sets or units with a peripheral position have the most distinct profiles from the average profile (Lebart & Salem, 1994; Poudat & Landragin, 2017; see also Blanchard & Patou, 2003 for more details).

(2) The analysis of specificities assessed the distribution of the considered units (e.g., orthographic forms, lemmas) in the different parts of the corpus. According to the hypergeometric model, a unit can be considered as specific of one part when its frequency in this part is superior (positive specificity) or inferior (negative specificity) to the expected theoretical frequency. This analysis is based on the total number of occurrences of the corpus, the total number of occurrences in the part, the frequency of the unit in the corpus, and the frequency of the unit in the part. In our analyses, a score equal or greater than 2 (i.e., the banality threshold) indicated an over-use of the unit, whereas a score equal or less than -2 indicated an under-use of the unit. A unit with a score between $]-2;2[$ was considered as banal. A unit that is banal in all parts of the corpus belongs to the basic vocabulary of the corpus (Lebart & Salem, 1994; Leblanc, 2015; Poudat & Landragin, 2017). This analysis was performed in TXM considering two parts (i.e., rhythmic regular sequences and textural sound sequences). In this configuration, the positive specificities of one part (e.g., regular rhythmic sequences) are the negative specificities of the other (e.g., textural sound sequences), and vice versa. As units, we considered the lemmas of content words. We complemented this analysis by further analysing the immediate context of the lemmas (i.e., words that are preceding or

following the target unit) that were positively specific in each part (Lebart & Salem, 1994; Leblanc, 2015).

Finally, we analysed the participants' judgments on the three subjective scales⁴. For each scale, participants' judgments for the two musical conditions were compared with two-sided paired *t*-tests. We reported effect sizes in terms of d_z or d using G*Power software (Faul, Erdfelder, Lang, & Buchner, 2007).

RESULTS

1. Analyses of the number of words and lemmas produced for each sequence type (Figure 1). Participants produced significantly more words after the textural sound sequences (mean = 33.65, SD = 16.41) than after the regular rhythmic sequences (mean = 26.91, SD = 17.50), $t_1(24) = -4.45$, $p < .001$, $d_z = 0.88$, $t_2(10) = -4.71$, $p < .001$, $d = 2.72$. More specifically, they produced significantly more content words after textural sound sequences (mean = 17.17, SD = 8.63) than after regular rhythmic sequences (mean = 13.79, SD = 9.45), $t_1(24) = -4.06$, $p < .001$, $d_z = 0.82$, $t_2(10) = -4.22$, $p = .002$, $d = 2.43$. This advantage for textural sound sequences was confirmed for each content category: nouns, $t_1(24) = -3.46$, $p = .002$, $d_z = 0.69$, $t_2(10) = -3.46$, $p = .006$, $d = 2.02$; verbs, $t_1(24) = -2.53$, $p = .019$, $d_z = 0.51$, $t_2(10) = -2.51$, $p = .031$, $d = 1.44$; adjectives, $t_1(24) = -3.95$, $p < .001$, $d_z = 0.79$, $t_2(10) = -3.56$, $p = .005$, $d = 2.04$; and adverbs, $t_1(24) = -2.39$, $p = .025$, $d_z = 0.48$, $t_2(10) = -2.76$, $p = .020$, $d = 1.59$.

The analyses on lemmas provided consistent results: participants produced significantly more lemmas after textural sound sequences (mean = 24.71, SD = 9.26) than after regular rhythmic sequences (mean = 19.94, SD = 10.47), $t_1(24) = -4.47$, $p < .001$, $d_z = 0.89$, $t_2(10) = -5.71$, $p < .001$, $d = 3.30$. Moreover, they produced significantly more content lemmas after textural sound sequences (mean = 13.95, SD = 5.90) than after regular rhythmic sequences (mean =

⁴ Note that because of a programming error, data of two participants were lost for emotion and interest scales and data of one participant were lost for the three scales.

11.28, $SD = 6.69$), $t_1(24) = -4.14$, $p < .001$, $d_z = 0.84$, $t_2(10) = -5.54$, $p < .001$, $d = 3.20$. Further analysing the produced lemmas of each grammatical type revealed that the difference between the two sequence types was significant for nouns, $t_1(24) = -3.40$, $p = .002$, $d_z = 0.68$, $t_2(10) = -3.60$, $p = .005$, $d = 2.08$; verbs, $t_1(24) = -2.65$, $p = .014$, $d_z = 0.53$, $t_2(10) = -2.56$, $p = .028$, $d = 1.48$; adjectives, $t_1(24) = -4.22$, $p < .001$, $d_z = 0.84$, $t_2(10) = -3.67$, $p = .004$, $d = 2.12$; and adverbs, $t_1(24) = -2.46$, $p = .022$, $d_z = 0.49$, $t_2(10) = -3.81$, $p = .003$, $d = 2.20$.

The generalized linear mixed model confirmed these results: the main effect of musical condition (Rhythmic vs. Textural sequences) was significant for content words, $X^2(1, N = 25) = 9.81$, $p = .002$, as well as for content lemmas, $X^2(1, N = 25) = 10.40$, $p = .001$.

Note that for 1.44 musical sequences on average, participants did not verbalize evocations (e.g. just saying “I don’t know” for example). More specifically, they did not verbalize for 1.16 rhythmic sequences and for 0.28 textural sequences on average.

[Figure 1 about here]

2. Analyses of the semantic content of the productions. (1) Factorial correspondence analysis: The first two axes explained respectively 16.6% and 10.8% of the total variance. The first axis discriminated between textural sound sequences on the left and regular rhythmic sequences on the right (see Figure 2, top)⁵. In addition, the horizontal distribution also showed two contrasting semantic fields, notably concepts relative to atmospheres, places and negative emotions on the extreme left (>-1) (e.g., castle, forest, thriller, stressful, black, gloomy) versus

⁵ Note that this opposition between textural sound sequences and regular rhythmic sequences was consistently observed regardless of the minimal frequency considered (from 2 to 10) to perform the analysis.

lemmas relative to movement, social behaviour and positive emotions on the extreme right (>1) (e.g., dance, rhythm, bar, move, celebrate, happy) (Figure 2, down).

[Figure 2 about here]

(2) Analysis of specificities: The analyses of specificities revealed several lemmas with a specificity score over banality threshold for each musical condition (see Table 1). Some specificity scores were higher for the textural sound sequences than for the regular rhythmic sequences. In contrast, more specific lemmas were observed for rhythmic sequences than for textural sequences (16 vs. 10). For some of these specific lemmas, their immediate context allowed us to get a more detailed picture of the content of the production, notably specifying its semantics: (a) for specific lemmas of rhythmic sequences: (1) “*rythmer*” [to punctuate] was mainly used in the past participle form “*rhythmé*”; (2) “*danse*” [dance] was notably associated with noun complements, for example: *dance class, dance area, street dances, dance studio*; (3) “*groupe*” [group] was mostly used with the meaning of group of musicians; (4) “*simplement*” [simply] was always used in the expression “*tout simplement*” [plainly]; (5) “*grand-chose*” [big thing] was always used with a negative form (e.g., “*pas grand-chose*” [not much]) to express the lack of things to say (e.g., “*ça ne m’a pas évoqué grand-chose*” [it didn’t mean much to me]). (b) for specific lemmas of textural sound sequences: (1) “*film*” [movie] was notably associated with noun complements and adjectives: mostly “*horror film*”, but also *crime movie, old movie, dramatic movie, sad movie, scary movie, suspense movie, black and white film*; (2) “*classique*” [classical] was associated with various nouns: mostly *classical music*, but also *classical dance, classical era, classical environment*; (3) “*noir*” [black] was used in the following expressions: *black and white film, black and grey film, black forest*; (4) “*scène*” [scene] was associated with various contexts:

movie scene, underwater scene, sad scene, suspense scene, crucial scene, scary scene, romantic scene.

[Table 1 about here]

3. Subjective scales judgments. Regular rhythmic sequences were judged as emotionally more positive (mean = 3,98, SD = 0.38) than the textural sound sequences (mean = 1.95, SD = 0.49), $t_1(20) = -16.74$, $p < .001$, $d_z = 3.61$, $t_2(10) = 15.12$, $p < .001$, $d = 8,61$. For the other two scales assessing interest and pleasure, the judgments did not significantly differ between the two musical conditions, all p -values $\geq .503$ for subject analyses, and all p -values $\geq .081$ for item analyses (Figure 3).

[Figure 3 about here]

DISCUSSION

The present study aimed to investigate the effect of two musical parameters (texture and rhythm) on semantic activation. In the textural sound sequences, the sound matter (produced by the fusion of many instruments into a slowly evolving texture) was the central element, without audible pulse or meter, whereas the regular rhythmic sequences contained a strong regular underlying beat and clear metrical structures, played with various musical instrument timbres. We used a semantic evocation task, which asked participants to describe what the music made them think about. We observed that participants produced more words and more lemmas after textural sound sequences than after regular rhythmic sequences. This pattern was observed when considering all grammatical categories together, and also in particular when considering those that reflect the lexical structure of a corpus (i.e., content unit: nouns,

verbs, adjectives, adverbs). The production of more content lemmas, and not only of more content words, after textural sound sequences suggests that participants activated indeed more concepts for these sequences. The obtained data thus suggest greater semantic network activation after textural sound sequences. However, it does not allow us to conclude about differences in the level of activation (i.e., weaker vs. strong activation). Moreover, based on Neely (1977), two mechanisms could operate in our task: spontaneous processing (i.e., automatic activation of a semantic content) and an intentional cognitive processing, which is controlled and directed (i.e., active research and elaboration of a semantic content). The sequence duration (i.e., 17 seconds), which was longer than the duration of sequences used in previous studies (e.g., 10 seconds in Koelsch et al., 2004, 1 second in Daltrozzo & Schön, 2009a, 2009b), might have been helpful for active research and elaboration. In addition, semantic activations might have occurred while listening to the music, but also when participants describe what the music made them think about⁶.

Our quantitative results thus confirm the efficiency of timbre to convey concepts (Daltrozzo & Schön, 2009a; Grieser Painter & Koelsch, 2011) and also suggest that when timbre is specifically manipulated in musical material, it promotes greater semantic network activation than does music oriented on rhythm and beat aspects. The analyses of the content of the productions, notably regarding the specific lemmas for rhythmic regular sequences, revealed the use of some timbre names (e.g., “drums”, “percussion” appeared as specific lemmas). These names highlight the role of musical dimensions (such as pitch height, loudness, and timbre) in music semantics, which give information about the source, in particular when the different timbres can be clearly separated in auditory streams (Schlenker, 2017). The musical timbres used in the regular rhythmic sequences can be more easily

⁶ Note that the experimenter was not blind to the hypotheses. However, even though we cannot exclude the influence of potential experimenter expectancy effects on the amount of productions, it seems difficult to explain how it would influence the content of the productions.

labelled, in particular as they are more distinct and familiar (e.g., bass drum, snare drum, cymbal). In contrast, for the textural sound sequences which correspond to blended timbres and sound sources, it was difficult to identify and to label the sound sources. This finding fits well within the framework proposed by Schön et al. (2009) in the context of conceptual priming paradigms, even though the authors presented it as a tentative model that needs to be clarified and developed in future studies. This model has been constructed on the basis of models of semantic processing (McNamara, 2005) and distributed network models (Anderson, 1993; McClelland & Rumelhart, 1986). It postulates an amodal concept lexicon, accessible via words as well as sounds. The amodal concept lexicon could thus be the link between concepts activated by sounds and concepts activated by words. When reading a word (as in the context of some of the conceptual priming experiments), three different levels specific to language processing would be successively activated: the letter level, the lexical level, and the semantic level. Finally, the amodal concept lexicon is activated. When listening to a sound, acoustic features are first extracted from the sound. Then, this activation can take two different routes, either it directly activates the amodal concept lexicon or it activates first the sound lexicon, which contains labels, or the emotional level before then reaching the amodal concept lexicon. A clearly identifiable environmental sound would activate directly a specific item in the sound lexicon (e.g., a barking sound should activate the label “dog”). In contrast, a sound difficult to identify would possibly not find a precise label in the sound lexicon. It would thus either activate directly representations in the amodal concept lexicon, or go to it via the emotional level. Once a concept is activated in the amodal concept lexicon, this activation would spread to the semantic level of the verbal part of the model. It allows explaining the previously reported conceptual priming effects between music and language (Schön et al., 2009). This model could also potentially be applied in the context of our experiment: The musical sequences might activate the sound lexicon, the emotional level and

the amodal concept lexicon. Then, activation of the semantic and lexical level might allow for the verbalization required by the task. In sum, we can make the hypothesis that textural sound sequences might have not found direct, appropriate entries in the sound lexicon and thus activated a larger set of concepts in the amodal concept lexicon. In contrast, regular rhythmic sequences rather activated labels (e.g., musical instrument names) in the sound lexicon. In other words, when sound sources are not or hardly identifiable, music could trigger more semantic elaboration and processes less strongly related to labelling. The framework proposed by Schön et al. (2009) thus provides interesting perspectives for future research along these lines, but – as illustrated by the discussion of our data here – it requires new experimentations directly designed to test and elaborate it.

The second contribution of our data is to provide some evidence that the content of the productions seems to be strongly influenced by the characteristics of the music. The factorial correspondence analysis clearly distinguished the lexical profiles of the productions for the textural sequences and the lexical profiles of the productions for the rhythmic sequences. The meaning of lemmas specifically associated with the regular rhythmic sequences was related to synchronized movement (e.g., “dance”) and social concepts (e.g., “party”, “group” (of musicians)), which is consistent with previous research showing that rhythm is related to movement and elicits motor synchronisation (e.g., Repp & Su, 2013; Su & Pöppel, 2012; Zentner & Eerola, 2010). Motor synchronisation is linked to social behaviour as moving together in synchrony increases cooperation and affiliation between individuals (Hove, 2009; Valdesolo, Ouyang & DeSteno, 2010; Wiltermuth & Heath, 2009). In a semantic priming paradigm with pictures, Zhou et al. (2015) reported an association between temporal musical features and the expression of movement. In a related vein, Taruffi, Pehrs, Skouras, & Koelsch, (2017) observed that while listening to happy and rhythmic music, thought content was predominantly characterized by dance imagery. For the textural sound sequences, the

specific lemmas were film-related terms (i.e., “movie” (dark movies), “suspense”, “scene”) and emotional terms, as discussed below.

The control task in our study revealed that participants judged the two sequence types as equally interesting and pleasant. Consequently, the increased semantic activation for textural sound sequences could not be due to different levels of interest or enjoyment (in comparison to the rhythmic sequences). The overall data pattern rather suggests that the two musical conditions might trigger different cognitive mechanisms, which explain the amount of words and lemmas produced. While textural sound sequences might trigger imaginative processes, rhythmic regular sequences seem to encourage labelling processes (i.e., sound sources and musical characteristics). Taken together, previous findings and our results suggest that (1) texture and timbral features are effective for the activation of semantic representations in long-term memory (see also Daltrozzo & Schön, 2009a; Grieser Painter & Koelsch, 2011), (2) rhythmic and metrical structures influence other linguistic processes, such as phonological processing (e.g., Cason & Schön, 2012) or syntax processing (in children: Bedoin, Brisseau, Molinier, Roch & Tillmann, 2016; Chern, Tillmann, Vaughan & Gordon, 2018; Przybylski et al., 2013; in adults: Canette et al., 2019, 2020b; Kotz, Gunter & Wonneberger, 2005). Listening to a regular rhythmic sequence would adjust internal oscillators and temporal attention (see Jones, 1976, 2019 for the Dynamic Attending Theory). As speech processing requires neural synchronization and temporal attention, it directly benefits from a pre-synchronization by music. The study by Canette et al., (2020a) conducted in children combined the same musical sequences as used here with two linguistic tasks, a semantic evocation task (as the one used in the present study) and a syntax processing task (i.e., grammaticality judgments, as used in previous rhythmic priming studies, e.g., Przybylski et al., 2013). The results showed that rhythm and texture had differential effects on the two tasks: rhythm enhanced grammaticality judgments compared to texture, and texture stimulated

more verbalization (i.e., more words and lemmas produced) than did rhythm. Our present results obtained in an adult population confirmed the enhanced effectiveness of texture (in comparison to rhythm) on semantic activation: Although verbal productions of adults were overall more numerous than those of children, they showed a similar data pattern than did the productions of children. The more detailed semantic analyses performed here (e.g., taken in consideration the context) provided further insights into the effect of the musical sequence on language processing.

The further control task reported in addition an effect on emotional valence. While the two sequences types did not differ on the subjective scales of interest and enjoyment, participants' judgments on the subjective scale of emotion indicated that the two sequence types differed in the evoked emotional valence: participants judged the textural sound sequences as sadder than the regular rhythmic sequences. Accordingly, analyses of participants' productions revealed that regular rhythmic sequences evoked words about positive emotions (e.g., "happy") whereas textural sound sequences evoked words about negative emotions (e.g., "sad", "frightening", "fear"). Although semantic activation by music can occur without the evocation of a strong emotional context (Grieser Painter & Koelsch, 2011; Zhou et al., 2015), emotion seems an important parameter for extra-musical meaning, as shown by affective priming studies using emotional words as targets (Sollberger et al., 2003; Steinbeis & Koelsch, 2011). These observations related to emotion can be discussed in relation to the model proposed by Schön et al. (2009): this model integrates an emotional level that precedes the amodal concept lexicon. Emotion also influence mind-wandering: a recent study has shown that the valence of music influence mind-wandering during music listening (Taruffi et al., 2017), notably with stronger mind-wandering for sad compared to happy music, independently of tempo. Specifically, during sad music listening, participants might direct their attention inwards, and thought content might mainly refer to emotions and

natural elements. Taruffi et al. (2017) suggest that music is able to trigger specific mental processes depending on the emotional valence of the music. Similarly, in our study, we observed that, among specific lemmas evoked for each sequence type, there were more lemmas about emotion for textural sound sequences than for regular rhythmic sequences.

Future studies could now test more precisely the contribution of each parameter (i.e., texture, rhythm, and emotional valence), even though disentangling the effect of the musical characteristics (i.e., texture, rhythm) from the effect of emotional valence might be challenging as structural features and emotion are entwined. For example, rhythm and tempo are key elements for mood induction in music: a dotted rhythm with a very slow tempo would evoke a sad emotion, whereas a syncopated rhythm with a fast tempo would rather evoke positive mood (about tempo and mode, see Peretz, 1998; Dalla Bella et al., 2001). Moreover, tonal variations are necessary for listeners' distinction between happy and sad musical excerpts, as also reflected in psychophysiological markers (Khalifa, Roy, Rainville, Dalla Bella & Peretz, 2008).

CONCLUSION

Our present study showed that specific musical parameters can modulate the extent of semantic activation. Our results revealed that sound texture promotes a greater semantic network activation than do rhythmic structures. These findings seem to be due to cognitive processes triggered by the musical parameters and the evoked emotions. They suggest that, for the investigation of the effect of music on linguistic processing, the control of these parameters as well as of perceived emotion is highly relevant. Furthermore, the findings are also interesting for the use of music as a therapeutic tool for training and rehabilitation of language disorders (e.g., Fiveash et al., submitted; Fujii & Wan, 2014; Schön & Tillmann, 2015).

ACKNOWLEDGEMENTS

We thank Marion Bendinelli, Virginie Lethier and Nicole Salzard (ELLIADD laboratory, University of Besançon, France) for their support with the software TXM. We also thank Anna Fiveash for her help with the generalized linear mixed model and the figures. The team "Auditory cognition and psychoacoustics" is part of the LabEx CeLyA ("Centre Lyonnais d'Acoustique", ANR-10-LABX-60) and L-HC is supported by a post-doctoral grant of the LabEx CeLyA.

REFERENCES

- Anderson, J. A. (1993). *The BSB Model: A simple nonlinear autoassociative neural network*. In M. Hassoun (Ed.), *Associative neural memories* (p. 77–103). New York: Oxford University Press.
- Audacity Team (2021). Audacity(R): Free Audio Editor and Recorder [Computer application]. Version 2.2.2 retrieved from <https://audacityteam.org/>
- Bedoin, N., Brisseau, L., Molinier, P., Roch, D., & Tillmann, B. (2016). Temporally regular musical primes facilitate subsequent syntax processing in children with Specific Language Impairment. *Frontiers in Neuroscience*, *10*, 245. <https://doi.org/10.3389/fnins.2016.00245>
- Blanchard, P., & Patou, C. (2003). Les usages de l'analyse factorielle dans les revues de sciences sociales. In Blanchard P., & Ribemont, T. (dir.). *Méthodes et outils des sciences sociales. Innovation et renouvellement*. L'Harmattan, collection Cahiers Politiques, Paris.
- Brooks M. E., Kristensen K., van Benthem K. J., Magnusson A., Berg C. W., Nielsen A., Skaug H. J., Maechler M., & Bolker B. M. (2017). glmmTMB Balances Speed and Flexibility Among Packages for Zero-inflated Generalized Linear Mixed Modeling. *The R Journal*, *9*(2), 378–400. .
- Canette, L-H., Bedoin, N., Lalitte, P., Bigand, E., & Tillmann, B. (2019). The regularity of rhythmic primes influences syntax processing in adults. *Auditory Perception and Cognition*, *2*(3), 163-179.
- Canette, L-H., Lalitte, P., Bedoin, N., Pineau, M., Bigand, E., & Tillmann, B. (2020a). Rhythmic and textural sound sequences differently influence syntax and semantic processing in children. *Journal of Experimental Child Psychology*, *191*. <https://doi.org/10.1016/j.jecp.2019.104711>

- Canette, L.-H., Fiveash, A., Krzonowski, J., Corneyllie, A., Lalitte, P., Thompson, D., Trainor, L., Bedoin, N., & Tillmann, B. (2020b). Regular rhythmic primes boost P600 in grammatical error processing in dyslexic adults and matched controls. *Neuropsychologia*, *138*. doi: 10.1016/j.neuropsychologia.2019.107324.
- Cason, N., & Schön, D. (2012). Rhythmic priming enhances the phonological processing of speech. *Neuropsychologia*, *50*(11), 2652-2658. <https://doi.org/10.1016/j.neuropsychologia.2012.07.018>
- Chern, A., Tillmann, B., Vaughan, C., & Gordon, R. L. (2018). New evidence of a rhythmic priming effect that enhances grammaticality judgments in children. *Journal of Experimental Child Psychology*, *173*, 371-379. <https://doi.org/10.1016/j.jecp.2018.04.007>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd edition). Hillsdale, NJ: Erlbaum.
- Dalla Bella S., Peretz, I., Rousseau, L., & Gosselin, N. (2001). A developmental study of the affective value of tempo and mode in music. *Cognition*, *80*, B1-B10.
- Daltrozzo, J., & Schön, D. (2009a). Conceptual processing in music as revealed by N400 effects on words and musical targets. *Journal of Cognitive Neuroscience*, *21*(10), 1882-1892. <https://doi.org/10.1162/jocn.2009.21113>
- Daltrozzo, J., & Schön, D. (2009b). Is conceptual processing in music automatic? An electrophysiological approach. *Brain Research*, *1270*, 88-94. <https://doi.org/10.1016/j.brainres.2009.03.019>
- Faul, F., Erdfelder, E., Lang, A-G., & Buchner, A. (2007). G*power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191.

- Fox, J., & Weisberg, S. (2011). *An {R} Companion to Applied Regression* (Second). SAGE Publications, Inc. <http://socserv.socsci.mcmaster.ca/jfox/Books/Companion>
- Fujii, S., & Wan, C. Y. (2014). The role of rhythm in speech and language rehabilitation: the SEP hypothesis. *Frontiers in Human Neuroscience*, 8, 777. <https://doi.org/10.3389/fnhum.2014.00777>
- Grieser Painter, J. G., & Koelsch, S. (2011). Can out-of-context musical sounds convey meaning? An ERP study on the processing of meaning in music. *Psychophysiology*, 48, 645-655. <https://doi.org/10.1111/j.1469-8986.2010.01134.x>
- Hansen, J., & Melzner, J. (2014). What you hear shapes how you think: Sound patterns change level of construal. *Journal of Experimental Social Psychology*, 54, 131-138. <https://doi.org/10.1016/j.jesp.2014.05.002>
- Heiden, S., Magué, J-P., & Pincemin, B. (2010). TXM : Une plateforme logicielle open-source pour la textométrie – conception et développement. Sergio Bolasco, Isabella Chiari, Luca Giuliano. 10th International Conference on the Statistical Analysis of Textual Data - JADT 2010, Jun 2010, Rome, Italie. Edizioni Universitarie di Lettere Economia Diritto, Statistical Analysis of Textual Data - Proceedings of 10th International Conference Journées d'Analyse statistique des Données Textuelles, 2(3), p.1021-1032, 2010.
- Hove, M. J., & Risen, J. L. (2009). It's all in the timing: interpersonal synchrony increases affiliation. *Social Cognition*, 27(6), 949-960. <https://doi.org/10.1521/soco.2009.27.6.949>
- Jones, M. R. (1976). Time, our lost dimension: toward a new theory of perception, attention and memory. *Psychological review*, 83(5), 323-355. <http://dx.doi.org/10.1037/0033-295X.83.5.323>
- Jones, M. R. (2019). *Time will tell: A theory of dynamic attending*. New York, NY: Oxford University Press.

- Khalifa, S., Roy, M., Rainville, P., Dalla Bella, S., & Peretz, I. (2008). Role of tempo entrainment in psychophysiological differentiation of happy and sad music? *International Journal of Psychophysiology*, 68(1), 17-26. <https://doi.org/10.1016/j.ijpsycho.2007.12.001>
- Koelsch, S. (2011). Towards a neural basis of processing musical semantics. *Physics of Life Reviews*, 8, 89-105. <https://doi.org/10.1016/j.plrev.2011.04.004>
- Koelsch, S., Kasper, E., Sammler, D., Schulze, K., Gunter, T., & Friederici, A. D. (2004). Music, language and meaning: brain signatures of semantic processing. *Nature Neuroscience*, 7(3), 302-307. <https://doi.org/10.1038/nn1197>
- Koelsch, S., & Siebel, W. A. (2005). Towards a neural basis of music perception. *Trends in Cognitive Sciences*, 9(12), 578-584. <https://doi.org/10.1016/j.tics.2005.10.001>
- Kotz, S. A., Gunter, T. C., & Wonneberger, S. (2005). The basal ganglia are receptive to rhythmic compensating during auditory syntactic processing: ERP patient data. *Brain and Language*, 95, 70-71. <https://doi.org/10.1016/j.bandl.2005.07.039>
- Lalitte, P. (2013). Texture et émergence perceptive en musique contemporaine. In F. Bort & V. Dupond (Eds.). *Texte, texture, textile. Variations sur le tissage dans la musique, les arts plastiques et la littérature*. Dijon, Éditions Universitaires de Dijon, 121-133.
- Lebart, L. & Salem, A. (1994). *Statistique Textuelle*. Paris : Dunod.
- Leblanc, J.-M. (2015). Proposition de protocole pour l'analyse des données textuelles : pour une démarche expérimentale en lexicométrie. *Nouvelles perspectives en sciences sociales*, 11(1), 25-63. <https://doi.org/10.7202/1035932ar>
- Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, 44(2), 314-324. <https://doi.org/10.3758/s13428-011-0168-7>

- McClelland, J. L., Rumelhart, D. E., & the PDP Research Group. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition* (Vol. II). Cambridge, MA: MIT Press.
- McNamara, T. P. (2005). *Semantic priming: Perspectives from memory and word recognition*. New York: Psychology Press.
- Meyer, L. (1956). *Emotion and Meaning in Music*. Leonard B. Meyer. Chicago, University of Chicago Press.
- Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General*, 106(3), 226-254. <http://dx.doi.org/10.1037/0096-3445.106.3.226>
- Patel, A. D. (2008). *Music, language and the brain*. New York: Oxford University Press.
- Peirce, C. S. (1931-58), *Collected Papers of Charles Sanders Peirce*, 8 vols., ed. C. Hartshome, P. Weiss, & A. W. Burks. Cambridge, Mass.: Harvard University Press.
- Peretz, I., Gagnon, L., & Bouchard, B. (1998). Music and emotion: perceptual determinants, immediacy, and isolation after brain damage. *Cognition*, 68(2), 111-114.
- Poudat, C., & Landragin, F. (2017). *Explorer un corpus textuel. Méthodes – pratiques – outils*. Louvain-la-Neuve, Belgique : De Boeck Supérieur.
- Poulin-Charronnat, B., Bock, B., Grieser, J., Meyer, K., & Koelsch, S. (2006). More about music, language and meaning: the follow-up of Koelsch et al. (2004). In: Baroni, M., Addessi, A.R., Caterina, R., Costa, M. (Eds.), *Proceedings of the 9th International Conference on Music Perception and Cognition (ICMPC9)*, Bologna/Italy, p. 1855. August 22–26 2006.

- Repp, B. H., & Su, Y-H. (2013). Sensorimotor synchronization: A review of recent research (2006-2012). *Psychonomic Bulletin & Review*, 20(3), 403-452. <https://doi.org/10.3758/s13423-012-0371-2>
- Przybylski, L., Bedoin, N., Krifi-Papoz, S., Herbillon, V., Roch, D., Léculier, L., Kotz, S., & Tillmann, B. (2013). Rhythmic auditory stimulation influences syntactic processing in children with developmental language disorders. *Neuropsychology*, 27(1), 121-131. <http://dx.doi.org/10.1037/a0031277>
- Rowell, L. (1983). *Thinking About Music: An Introduction to the Philosophy of Music*, Amherst: The University of Massachusetts Press.
- RStudio Team (2016). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA URL: <http://www.rstudio.com/>.
- Schlenker, P. (2017). Outline of music semantics. *Music Perception*, 35(1), 3-37. <https://doi.org/10.1525/mp.2017.35.1.3>
- Schlenker, P. (2019). Prolegomena to music semantics. *Review in Philosophy & Psychology*, 10, 35-111. <https://doi.org/10.1007/s13164-018-0384-5>
- Schmid, H. (1994). *Probabilistic Part-of-Speech Tagging Using Decision Trees*. In Proceedings of the Conference on New Methods in Language Processing, Manchester, UK, 1994.
- Schön, D., & Tillmann, B. (2015). Short- and long-term rhythmic interventions: Perspectives for language rehabilitation. *Annals of The New York Academy of Sciences*, 1337, 32-39. <https://doi.org/10.1111/nyas.12635>
- Schön, D., Ystad, S., Kronland-Martinet, R., & Besson, M. (2009). The evocative power of sounds: conceptual priming between words and nonverbal sounds. *Journal of Cognitive Neuroscience*, 22(5), 1026-1035. <https://doi.org/10.1162/jocn.2009.21302>
- Smith-brindle, R. (1966). *Serial Composition*, Oxford: Oxford University Press.

- Sollberger, B., Reber, R., & Eckstein, D. (2003). Musical chords as affective priming context in a word-evaluation task. *Music Perception, 20*(3), 263-282. <https://doi.org/10.1525/mp.2003.20.3.263>
- Steinbeis, N., & Koelsch, S. (2011). Affective priming effects of musical sounds on the processing of word meaning. *Journal of Cognitive Neuroscience, 23*(3), 604-621. <https://doi.org/10.1162/jocn.2009.21383>
- Su, Y-H., & Pöppel, E. (2012). Body movement enhances the extraction of temporal structures in auditory sequences. *Psychological Research, 76*(3), 373-382. <https://doi.org/10.1007/s00426-011-0346-3>
- Taruffi, L., Pehrs, C., Skouras, S., & Koelsch, S. (2017). Effects of sad and happy music on mind-wandering and the default mode network. *Scientific Reports, 7*, 14396. <https://doi.org/10.1038/s41598-017-14849-0>
- Valdesolo, P., Ouyang, J., & DeSteno, D. (2010). The rhythm of joint action: Synchrony promotes cooperative ability. *Journal of Experimental Social Psychology, 46*(4), 693-695. <https://doi.org/10.1016/j.jesp.2010.03.004>
- Wiltermuth, S. S., & Heath C. (2009). Synchrony and cooperation. *Psychological Science, 20*, 1. <https://doi.org/10.1111/j.1467-9280.2008.02253.x>
- Zentner, M., & Eerola, T. (2010). Rhythmic engagement with music in infancy. *PNAS, 107*(13), 5768-5773. <https://doi.org/10.1073/pnas.1000121107>
- Zhou, L., Jiang, C., Yingying, W., & Yang, Y. (2015). Conveying the concept of movement in music: An event-related brain potential study. *Neuropsychologia, 77*, 128-136. <https://doi.org/10.1016/j.neuropsychologia.2015.07.029>

APPENDIX 1

1. Textural sound sequences

T1: György Ligeti – *Chamber Concerto* for 13 instruments (1970), 2nd movement (from 0.36). The texture consists of a micropolyphony with superimposed chromatic lines with delayed attacks (Fl., Ob. d'am., Cl., Hn., Trb., Org. strings).

T2: Peter Eötvös – *Chinese Opera* for Ensemble (1986), scene 1 (from 4.20). The texture consists of sustained microtonals chords (2 Fl., 2 Ob., 3 Cl., 2 Bn., 2 Picc. Trp., 1 synthesizer DX7 Yamaha, strings).

T3: György Ligeti – *Lontano* for Orchestra (1967), from 0.19. The texture consists of a canon (all instruments play the same pitch series) over long durations in the high register (4 Fl., Ob., 4 Cl., 3 Bn., 3 Hn., Trp., Va., Vc.).

T4: György Ligeti – *Melodien* for Orchestra (1971), from 2.50. The texture consists of polyphony of sustained sounds in the medium-high register (Fl., Ob., Cl., Bn., Hn., Trp., Trb., strings)

T5: Tristan Murail – *Le lac* for Ensemble (2001), from 5.06. The texture consists of a few notes held in the high register, with vibrato variations, and a noisy sound maintained by the percussion (2 Fl., perc., Vl., Va., 2 Vc.).

T6: Roger Reynolds – *Symphony [Myths]* for Orchestra (1990), 1st movement, from 2.55, The texture consists of polyphony of sustained sounds in the high register with variations of dynamics (2 Picc., Fl., Vn. 1 & 2).

Fl. = Flute, *Ob.* = Oboe, *Picc.* = Piccolo, *Ob. d'am.* = Oboe d'amore, *Cl.* = Clarinette, *Bn.* = Bassoon, *Hn.* = Horn, *Trp.* = Trumpet, *Trb.* = Trombone, *Perc.* = percussion, *Vn.* = violin, *Va.* = Viola, *Vc.* = Cello, *Org.* = Organ

2. Regular rhythmic sequences

R1: Hi-hat (height notes) Snare drum (quarter notes, upbeat), bass drum (quarter, eighth, and sixteenth notes), claves (eighth notes in backbeats)

R2: Tambourine (height notes), Bongos (eighth notes), Hand Claps (quarter and eighth notes), Snare drum (quarter and eighth notes), Bass drum (half, dotted quarter, and eighth notes)

R3: Bongos (height notes), Hihat (eighth notes upbeat), Handclap (quarter notes, eighth notes), Synthesized bass sound (dotted quarter note, eighth note tied to half note)

R4: Ride cymbal (eighth notes), Bongos (eighth notes upbeat), Claves (eighth notes), Synthesized bass sound (quarter notes)

R5: Snare drum (quarter, eighth, and sixteenth notes), Wood-block (eighth notes), Synthesized bass sound (quarter notes, down beat)

R6: Maracas (eighth notes), Snare drum (quarter notes, eighth notes), Handclap (eighth notes upbeat), Synthesized bass sound (dotted quarter note, eighth note tied to half note)

APPENDIX 2

French version of the factorial correspondence analysis representing the lemmas (Figure 2, down)



SUPPLEMENTAL MATERIAL 1

The supplemental material can be found online at [*to be completed later*].

SUPPLEMENTAL MATERIAL 2

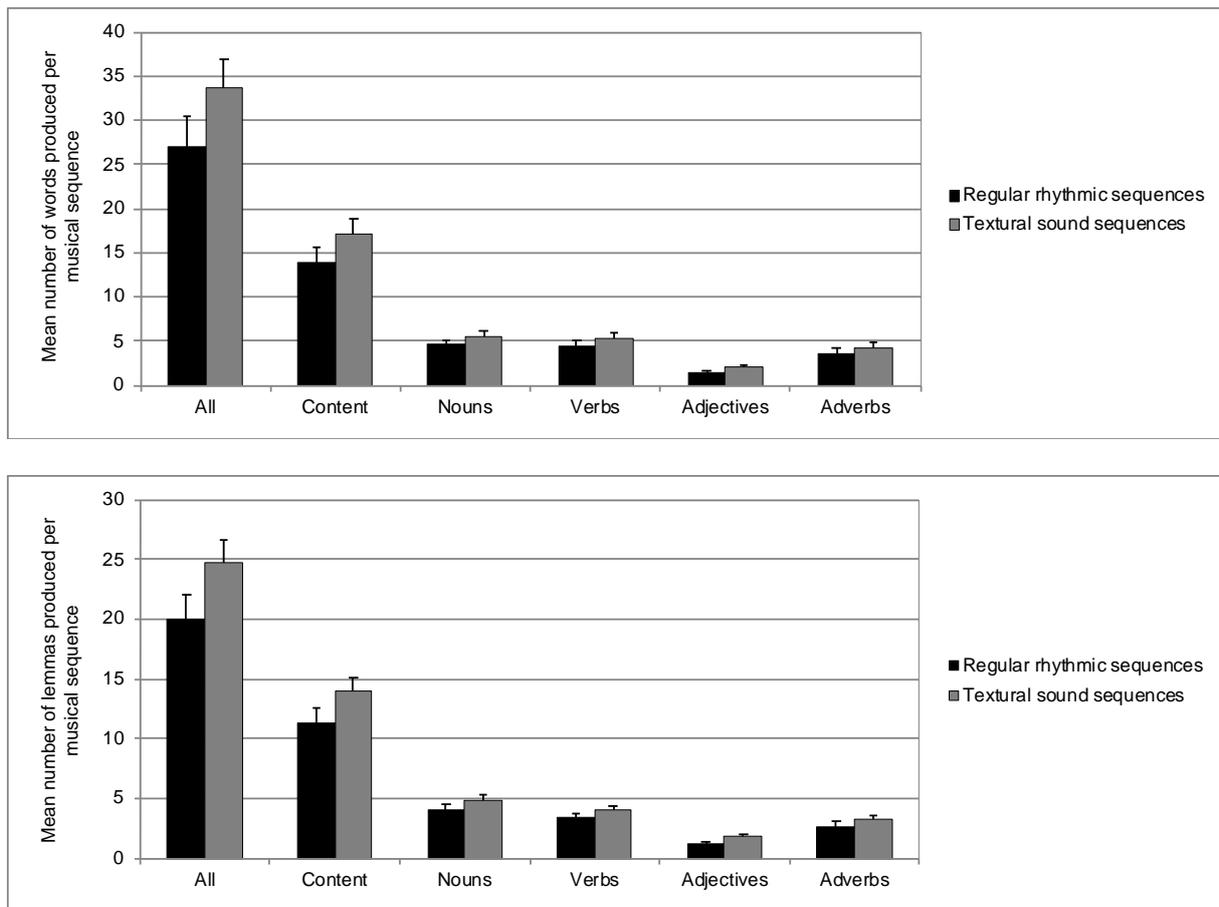


Figure 1. Mean number of words (Top) and lemmas (Bottom) produced per musical sequence, presented as function of the sequences (regular rhythmic sequences vs. textural sound sequences) for all productions, content words only, and separated for each subgroup: nouns, verbs, adverbs and adjectives. Error bars represent standard errors.

AUTHOR NOTE

Laure-Hélène Canette

Laure-Helene.Canette@u-bourgogne.fr

LEAD - CNRS - Université Bourgogne Franche-Comté - UMR 5022

Institut Marey - I3M

64 rue de Sully

21000 Dijon

France

TABLES

Table 1. Positive specific lemmas (i.e., score > 2) and their score of specificity for each musical condition.

Regular rhythmic sequences		Textural sound sequences	
danser [to dance]	4.9	film [movie]	9.7
fête [party]	3.9	horreur [horror]	7.5
joyeux [happy]	3.9	peur [fear]	4.6
rythmer [to punctuate]	3.9	triste [sad]	3.3
rythme [rhythm]	3.5	suspense [suspense]	3.0
batterie [drums]	2.8	angoissant [frightening]	2.8
danse [dance]	2.7	classique [classical]	2.1
groupe [group]	2.7	dramatique [dramatic]	2.1
percussion [percussion]	2.5	noir [black]	2.1
plage [beach]	2.5	scène [scene]	2.0
bar [bar]	2.1		
boum [bang]	2.1		
festival [festival]	2.1		
part (à part) [beyond]	2.1		
simplement [simply]	2.1		
grand-chose [big thing]	2.1		

FIGURE CAPTIONS

Figure 1. Boxplots for the mean number of content words (left side) and content lemmas (right side) produced per musical sequence, presented as function of the sequences (regular rhythmic sequences vs. textural sound sequences). Each participant is represented by a dot in each musical condition, and the two dots are connected. (See Supplemental Material 2 for the figures representing data as a function of grammatical category).

Figure 2. Projection of the first two factors of the factorial correspondence analysis, representing the 12 musical sequences (top) or the lemmas (bottom). See Appendix 2 for a French version of this Figure. The « XXX » refers to words that were not understandable in the recordings.

Figure 3. Boxplots for the mean results (judgments from 1 to 5) for regular rhythmic sequences and textural sound sequences on the three scales (the first scale assessed emotion, from “very sad” (1) to “very happy” (5); the second scale assessed interest, from “very boring” to “very interesting”; the third scale assessed pleasure, from “I don’t like at all” to “I like very much”). Each participant is represented by a dot in each musical condition, and the two dots are connected. Note that data of two participants were missing for emotion and interest scales and data of one participant was missing for the three scales (see method section).

FIGURES

Figure 1.

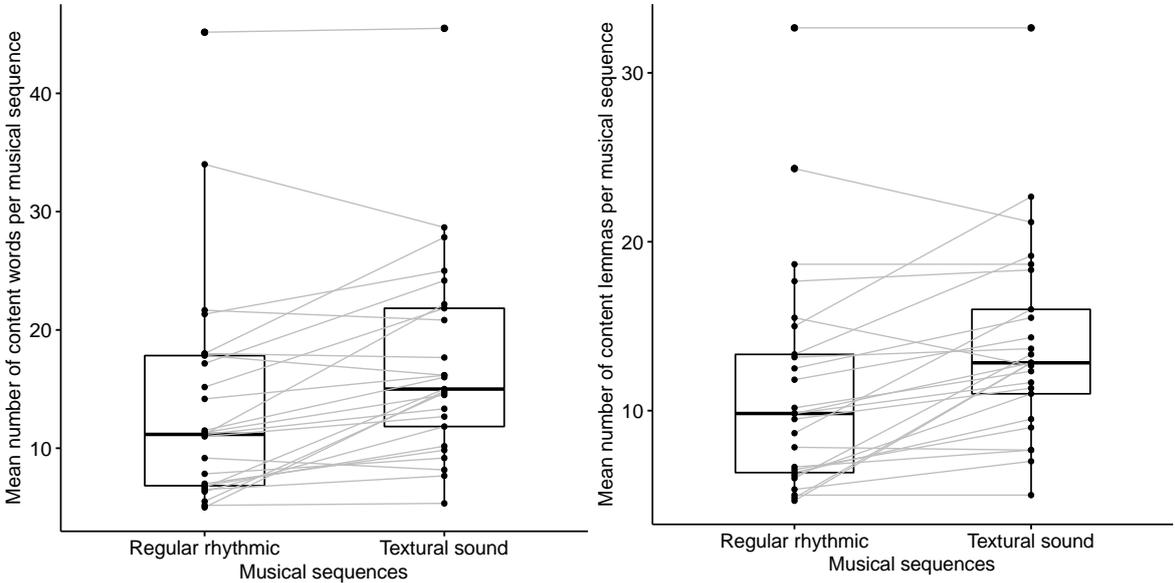
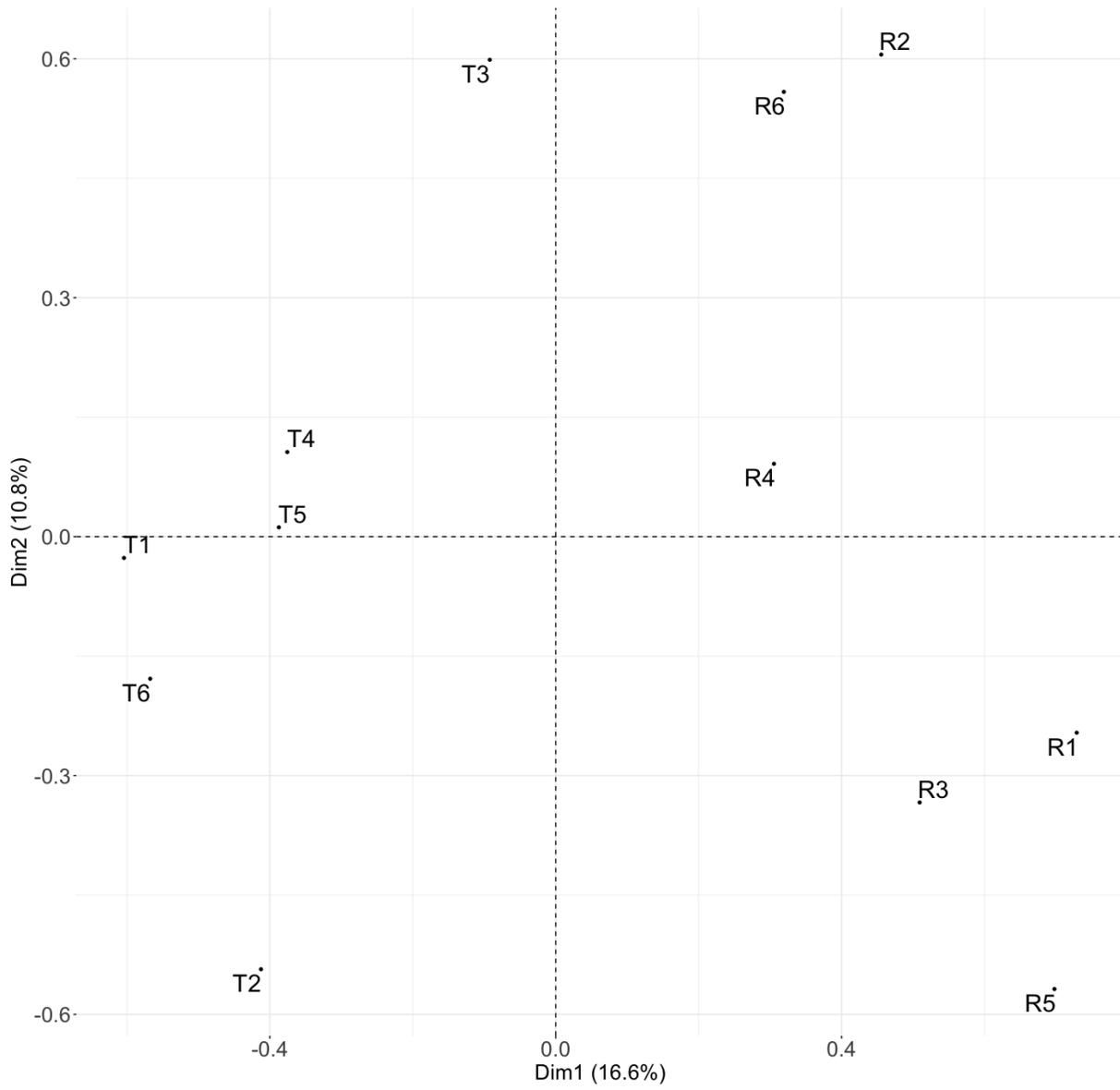


Figure 2.



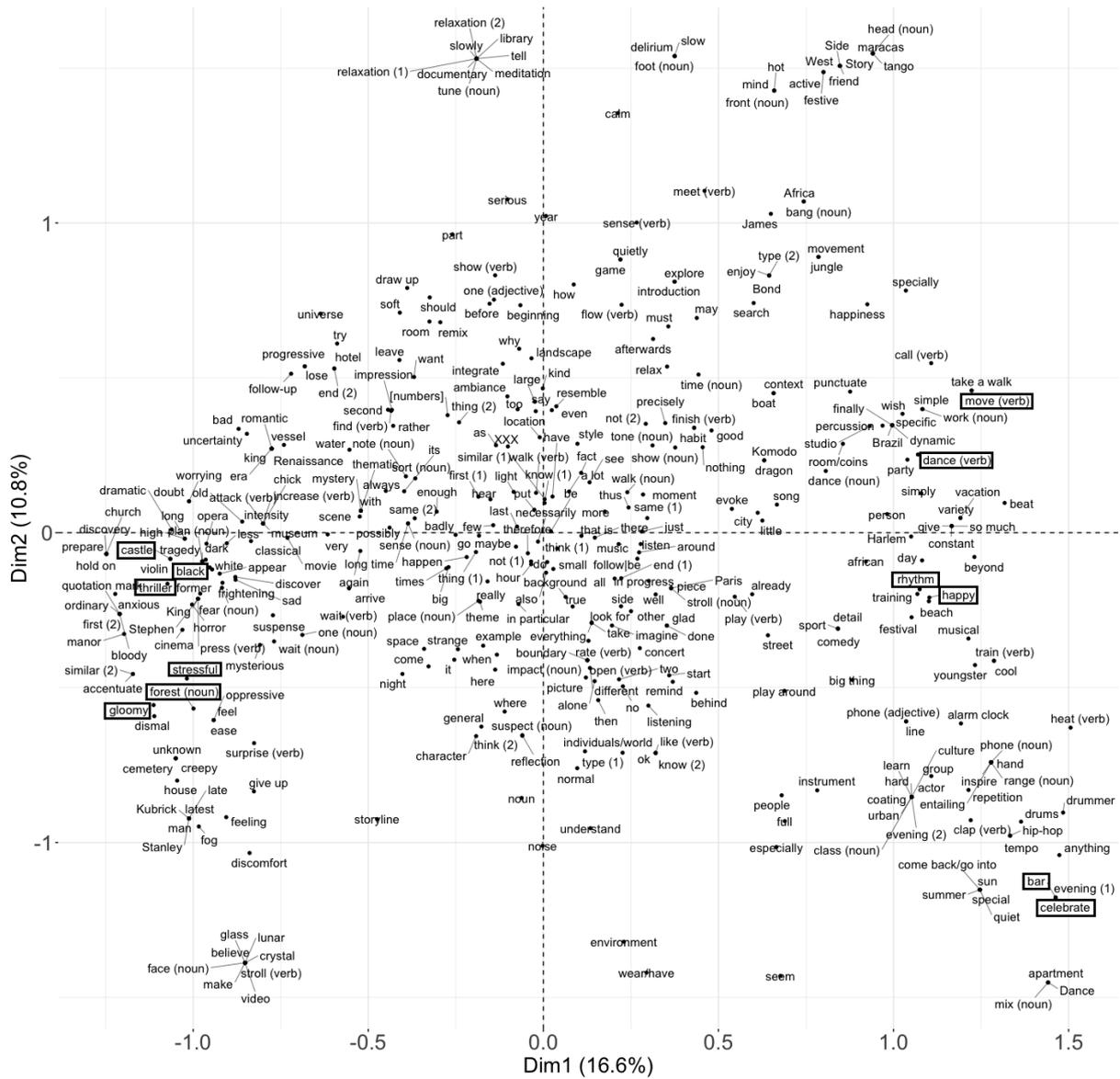


Figure 3.

