Cognition 107 (2008) 1070–1083

COGNITION

Brief article

The tonal function of a task-irrelevant chord modulates speed of visual processing

N. Escoffier [a,b,c], B. Tillmann [a,*]

[a] *Université Claude Bernard Lyon 1, Neurosciences Sensorielles, Comportement, Cognition, CNRS-UMR 5020, IFR 19, 50 Avenue Tony Garnier, F-69366 Lyon Cedex 07, France*
[b] *University of Georgia, Department of Psychology, Athens, GA 30602, USA*
[c] *National University of Singapore, Department of Psychology, Faculty of Arts and Social Sciences, 9 Arts Link, Singapore 117570, Singapore*

Received 7 July 2006; revised 15 October 2007; accepted 19 October 2007

Abstract

Harmonic priming studies have provided evidence that musical expectations influence sung phoneme monitoring, with facilitated processing for phonemes sung on tonally related (expected) chords in comparison to less-related (less-expected) chords [Bigand, Tillmann, Poulin, D'Adamo, and Madurell (2001). The effect of harmonic context on phoneme monitoring in vocal music. *Cognition, 81,* B11–B20]. This tonal relatedness effect has suggested two interpretations: (a) processing of music and language interact at some level of processing; and (b) tonal functions of chords influence task performance via listeners' attention. Our study investigated these hypotheses by exploring whether the effect of tonal relatedness extends to the processing of visually presented syllables (Experiments 1 and 2) and geometric forms (Experiments 3 and 4). For Experiments 1–4, visual target identification was faster when the musical background fulfilled listeners' expectations (i.e., a related chord was played simultaneously). In Experiment 4, the addition of a baseline condition (i.e., without an established tonal center) further showed that the observed difference was due to a facilitation linked to the related chord and not to an inhibition or disruption caused by the less-related chord. This outcome suggests the influence of musical structures on attentional mechanisms and that these mechanisms are

* Corresponding author. Tel.: +33 (0) 4 37 28 74 93; fax: +33 (0) 4 37 28 76 01.
E-mail address: barbara.tillmann@olfac.univ-lyon1.fr (B. Tillmann).

0010-0277/$ - see front matter
doi:10.1016/j.cognition.2007.10.007

shared between auditory and visual modalities. The implications for research investigating neural correlates shared by music and language processing are discussed.

Keywords: Musical priming; Audiovisual interaction; Attention; Music and language

1. Introduction

For language and music, context effects due to perceivers' knowledge about these structured systems have been reported. For language, target word processing is facilitated when preceded by a semantically related prime context (McNamara, 2005). For music, target chord processing is facilitated when preceded by a tonally related prime context (Bigand & Poulin-Charronnat, in press; Tillmann, Bharucha, & Bigand, 2000). Research in cognition and neuroscience has shown a growing interest in the study of the independence or dependence of language and music processing. Cognitive and neural resources that are either domain-specific or shared by language and music are under investigation (Besson, Faïta, Peretz, Bonnel, & Requin, 1998; Patel, Gibson, Ratner, Besson, & Holcomb, 1998, 2003; Peretz & Coltheart, 2003). For example, behavioral and neurophysiological data on the perception of sung music support either independence (e.g., Besson et al., 1998; Bonnel, Faïta, Peretz, & Besson, 2001) or dependence (e.g., Bigand, Tillmann, Poulin, D'Adamo, & Madurell, 2001; Samson & Zatorre, 1991) in processing.

Behavioral studies showed an influence of musical structures on the processing of sung phonemes (Bigand, Tillmann, Poulin, D'Adamo, & Madurell, 2001) and words (Poulin-Charronnat, Bigand, Madurell, & Peereman, 2005). In Bigand, Tillmann, Poulin, D'Adamo, and Madurell (2001), eight-chord sequences were sung on artificial syllables and participants performed speeded phoneme identification on the last chord (i.e., sung on /di/ or /du/). The last chord (i.e., target) was tonally related or less-related (i.e., functioning either as tonic or subdominant)[1] and was supposed to be expected or less-expected in the musical context. Phoneme identification was faster when the phoneme was sung on a related chord than on a less-related chord. This influence of musical structures extends to the processing of semantic relatedness for sung words in a lexical decision task (Poulin-Charronnat et al., 2005). Since phoneme-identification and lexical decision can be performed by focusing on linguistic information only, the musical relatedness effect suggests that musical structure processing is automatic (see also Justus & Bharucha, 2001) and musical expectations influence linguistic processing. These findings suggest that processing of musical structures and processing of both phonemes and words (as well as their semantic

[1] A critical feature of the Western tonal musical system is that tones and chords differ in their tonal function depending on the context (and its instilled key) in which they occur. The tonic chord is at the top of the hierarchy and is the most referential member of a key. It commonly serves as the final event in a musical phrase, eliciting a feeling of closure and completion. It is followed in importance by the dominant chord, the subdominant chord, other in-key chords and finally out-of-key chords.

relatedness to the context) interact at some level of processing. In addition to integrating these findings into the debate of dependence and independence of music and language processing, Bigand, Tillmann, Poulin, D'Adamo, and Madurell (2001) and Poulin-Charronnat et al. (2005) proposed an alternative hypothesis: The musical relatedness effect might be non-specific to language and be mediated by general attentional processes. In music, rhythmic, melodic and tonal-harmonic accents function together creating the dynamic shape of the musical sequence and, notably, modulating attention over time (Jones, 1987). For the sung chord sequences, the authors proposed that the related chords (i.e., tonally important events at the end of the sequence) function as harmonic accents and are "likely to work as (culturally based) attentional markers, which capture more attentional resources" (page B57, Poulin-Charronnat et al., 2005). The increased attentional resources in the related condition thus facilitate the processing of phonemes and semantic information in sung music (Bigand et al., 2001; Poulin-Charronnat et al., 2005).[2]

Our study investigated the interaction between music and language processing and the alternative, attentional hypothesis with a cross-modal priming paradigm. In the sung sequences, the to-be-processed linguistic information and musical structures were embedded in the same acoustic signal. In our study, sequences of visual syllables were presented while musical sequences, ending on related or less-related chords, were presented in synchrony. Experiments 1 and 2 investigated whether the interaction extends to the processing of dissociated, visually presented syllables. Experiments 3 and 4 investigated whether the influence of musical structures on event processing is restricted to language or extends to the processing of other visual material, notably geometric forms. Experiment 4 further added a musically neutral baseline condition (i.e., without installing a tonal center) to investigate whether the influence of tonal function on visual processing was due to facilitation and/or inhibition.

2. Experiment 1

2.1. Method

2.1.1. Participants

Sixteen University of Lyon students participated in Experiment 1. Number of years of musical training, as measured by instrumental instruction, ranged from 0 to 9, with a mean of 1.9 (standard deviation, SD = 1.0) and a median of 2.7.

[2] Interestingly, Jones' framework of dynamic attention has been applied not only for the modulation of attention in music, but also to the modulation of attention in speech (Cutler, 1976; Jones & Boltz, 1989; Port, 2003). For example, facilitated phoneme-monitoring for regular inter-stress timing in spoken word lists has been interpreted in terms of attentional rhythms, with regular speech timing improving speech communication (Quené & Port, 2005).

2.1.2. Material

The auditory stimuli were 12 eight-chord sequences from Bigand et al. (2001). A sequence was composed of six chords completed by two ending chords. These endings were constructed in such a way that the last chord functioned as either a tonic (related condition) or a subdominant (less-related condition). Sequences were generated with Cubase 5.1, Halion software and acoustic piano samples (The Grand). Each chord was played for 625 ms.

The visual stimuli were sequences of eight syllables presented in synchrony with the chords (see Fig. 1). On the basis of five consonants and five vowels, 22 CV-syllables were constructed and assigned in various combinations to the chord sequences (e.g., *da-fe-ku-po-fa-to-ke*), the last syllable being *di* or *du* (as in Bigand et al., 2001). Syllables were displayed in white at the center of a computer screen on a black background. Each of the first seven syllables was displayed for 380 ms with an inter-syllable-interval of 245 ms. The last syllable (the target) was displayed for 700 ms. The experiment was run with PsyScope software (Cohen, MacWhinney, Flatt, & Provost, 1993).

2.1.3. Procedure

Participants were instructed that sequences of syllables were displayed on the screen, and were asked to judge as quickly and as accurately as possible whether the last syllable of each sequence was *di* or *du*, by pressing one of two keys on the

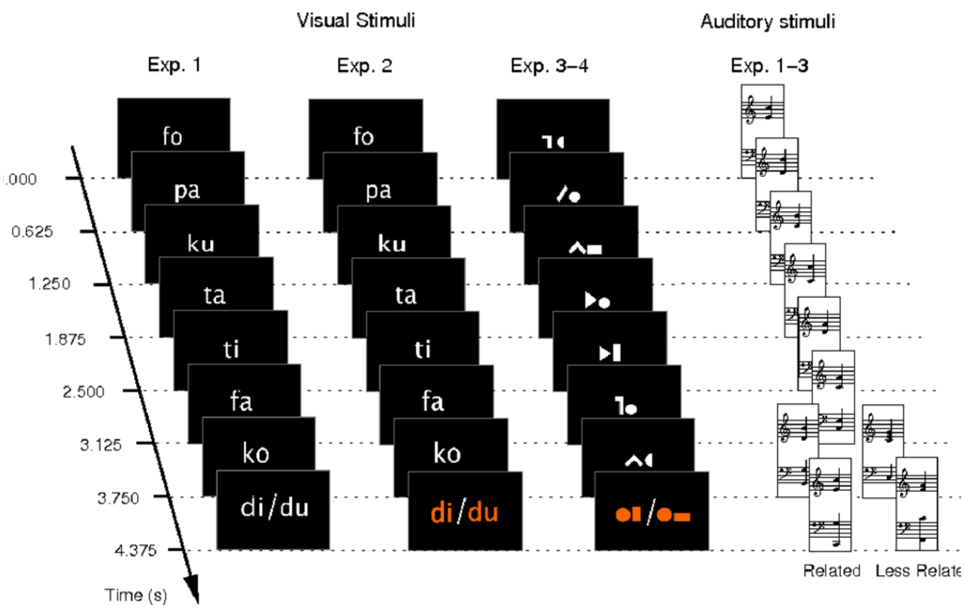

Fig. 1. Stimuli sequences and timing in the visual and auditory modalities. On the right, one example of musical sequences; the first six chords being identical and the ending chord, played in synchrony with the target, being either related (a *tonic* chord, noted I) or less-related to the prime (a *subdominant* chord, noted IV).

computer keyboard. Participants were informed that the syllable sequences were presented with background music, but their task concerned only the syllables. Incorrect responses were followed by an auditory feedback signal and a 250 ms noise burst indicated the trial's end. To start the next trial, participants pressed a third key. After four training sequences, 48 trials (i.e., the six related and six less-related musical sequences associated twice with a syllable sequence ending on *di* and on *du*) were presented in random orders.

2.2. Results

Accuracy was high overall (97%, Table 1) and correct response times (Fig. 2a) were analyzed with two 2×2 ANOVAs with Musical Relatedness (related/less-related) and Target syllable (*di*/*du*) as within-participants factors and either participants (F_1) or sequences (F_2) as random variable. The main effect of Musical relatedness was significant ($F_1(1,15) = 5.13$, $p < .05$, $MSE = 1679.19$; $F_2(1,5) = 115.48$, $p < .001$). Participants were faster at identifying the target syllable when the chord played in synchrony was musically related. The item analyses showed a significant effect of Target syllable, $F_2(1,5) = 13.44$, $p < .05$. No other effects were significant.

2.3. Discussion

Experiment 1 extended the musical relatedness effect observed by Bigand et al. (2001) for sung syllables to visually presented syllables. This outcome suggests that the interaction between language and music processing extends from the auditory to the visual modality. However, it might be argued that participants used the last chord's tonal function (i.e., tonic or subdominant) as a cue for the last syllable. In Western tonal music, the tonic serves as the terminal event in a musical phrase and elicits a feeling of closure and completion. In the related condition, the musical function of the tonic might have signaled the last (closing) item of the sequence and thus the to-be-judged target syllable. In the less-related condition, the subdominant, untypical at the end of musical phrases and inducing a feeling of incompleteness (Bigand & Pineau, 1997; Boltz, 1989), might have provided conflicting information, thus slowing down the identification of the last syllable as the target. To rule out this interpretation, Experiment 2 added a visual marker (i.e., color change) indicating the target syllable.

Table 1
Percentages of correct responses for Experiments 1–3 presented as a function of Target type and related and less-related conditions

Experiment 1 Harmonic relatedness			Experiment 2 Harmonic relatedness			Experiment 3 Harmonic relatedness		
Target	Related	Less related	Target	Related	Less related	Target	Related	Less related
di	96.4 (1.0)	97.4 (1.0)	*di*	95.5 (1.2)	94.9 (1.2)	Pair A	94.4 (1.3)	97.2 (1.3)
du	96.4 (1.0)	98.4 (1.0)	*du*	99.4 (1.2)	94.2 (1.2)	Pair B	95.6 (1.3)	92.9 (1.3)

Within-participant standard error of the mean (Masson & Loftus, 2003) are indicated in parenthesis.

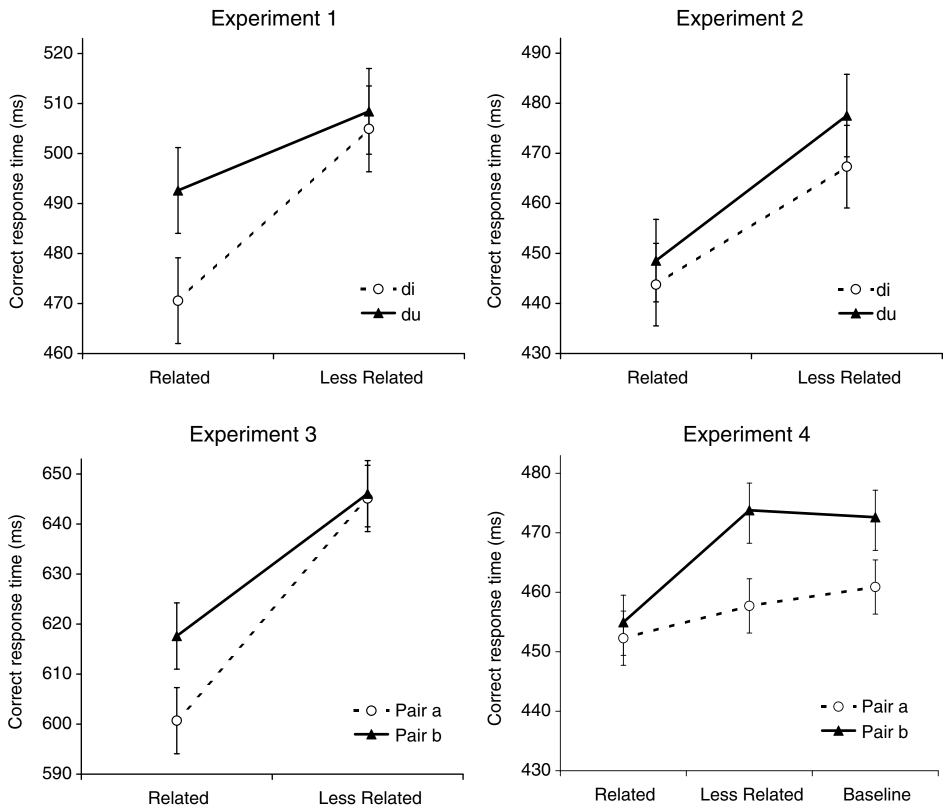

Fig. 2. Correct response times for visual target processing presented as a function of target type (*di*/*du* or Pair A/Pair B) and of musical relatedness (related and less-related conditions) for Experiments 1–3, and as a function of target pair and related, less-related and baseline conditions for Experiment 4. Error bars indicate within-participant standard error of the mean (Masson & Loftus, 2003).

3. Experiment 2

3.1. Method

3.1.1. Participants

Thirteen University of Georgia students participated in Experiment 2. Number of years of instrumental instruction ranged from 0 to 8, with a mean of 2.6 (SD = 3.1) and a median of 1.0.

3.1.2. Material and procedure

Auditory stimuli, visual stimuli and procedure were as described in Experiment 1, except that target syllables were displayed in orange and participants were explicitly told so.

3.2. Results

Accuracy was high overall (96%, Table 1). The ANOVAs of correct response times (Fig. 2b) only revealed a main effect of Musical relatedness ($F_1(1,12) = 11.49$, $p < .01$, $MSE = 781.1$; $F_2(1,5) = 5.29$, $p = .07$): Participants identified the target syllable faster when a related chord was played in the background.

3.3. Discussion

The color cue of the target syllable did not modify the response time pattern: Harmonic structures of the background music were processed and influenced syllable identification, even with attention focused on the visual material. This result can be interpreted both in terms of interaction between language and music processing and the alternative, attentional hypothesis, as previously proposed by Bigand et al. (2001) and Poulin-Charronnat et al. (2005) for sung music. The attentional effect of musical structures should be non-specific to language and should also apply to non-verbal information. In Experiment 3, a visual identification task of geometric shapes was performed with the same background music as in Experiments 1 and 2. While the hypothesis of language-specific interaction predicts no influence of musical structures, the attentional hypothesis predicts an effect of musical relatedness on shape identification, with facilitated processing for shapes presented in synchrony with related chords.

4. Experiment 3

4.1. Method

4.1.1. Participants

Twenty-one University of Lyon students participated in Experiment 3. None had participated in Experiment 1. Number of years of instrumental instruction ranged from 0 to 6, with a mean of 1.2 (SD = 1.9) and a median of 0.

4.1.2. Material

Auditory stimuli were as described in Experiment 1. To match the complexity and variety of syllables in Experiments 1 and 2, 10 geometric shapes were associated in pairs and the target pair was a circle with a vertical or a horizontal bar (Fig. 1). Pairs of shapes were presented in sequences of eight, synchronized to the chords. They were displayed at the center of the screen on a black background, with the first seven pairs in white and the target pair in orange.

4.1.3. Procedure

The procedure was as described in Experiment 2, except that participants were asked to judge as quickly and as accurately as possible whether the last pair of shapes (displayed in orange) was Pair A or Pair B.

4.2. Results

Accuracy was high overall (95%, Table 1). Correct response times (Fig. 2c), analyzed with Musical Relatedness and Target pair as within-participant factors, revealed only a main effect of Musical relatedness ($F_1(1,20) = 24.36$, $MSE = 1030.03$, $p < .001$); $F_2(1,5) = 13.35$, $p < .05$): target shapes were identified faster when presented together with a related chord.

4.3. Discussion

Experiment 3 extended the influence of musical relatedness on visual processing from syllables to geometric shapes. This finding suggests that the cross-modal effect of musical relatedness is not limited to language processing, but involves more general attentional mechanisms, which also affect non-verbal processing. These results could be integrated into Jones' (1987) attentional framework, as proposed previously, or the slowed-down processing of the visual information might be interpreted as reflecting a perturbation of the attentional system due to a disruptive effect of the less-expected subdominant chord. This interpretation is comparable to low-level oddball-like detection mechanisms, where the strength of the oddball's deviance from the standard modulates response times in detection tasks and the latency of the evoked MMN (Tiitinen, May, Reinikainen, & Näätänen, 1994). Our musical material construction can exclude a sensory-based oddball explanation for the observed difference (i.e., neither tonic nor subdominant occurred in the sequence prior to the last chord, see Bigand et al., 2001 and 2003 for a discussion). However, it might be argued that – because of the musical expectancy violation – the less-related chord reorients attention away from the visual information to the auditory information and thus slows down the visual task performance.

To test this alternative hypothesis, the cost and benefit patterns involved in the here observed audio–visual interaction were investigated in Experiment 4 by adding a musically neutral sequence (i.e., ending on a chord without a specific tonal function). The use of neutral baseline conditions (i.e., "neutral" in the sense of being neither related nor unrelated) have been used extensively in psycholinguistic research (e.g., Neely, 1976) and more recently in music perception research (e.g., Tillmann, Janata, Birk, & Bharucha, 2003). Facilitation and inhibition are defined as faster or slower response times relative to these neutral prime conditions. Musical baseline sequences are constructed in such a way that the prime context does not instill a tonal center and thus does not attribute a tonal function to the final chord. Previous experiments on music perception (i.e., auditory signal only) have revealed a cost of processing (i.e., inhibition) for less-related chords, but a benefit of processing (i.e., facilitation) for related chords (Tillmann et al., 2003). If the cross-modal effect is caused by a disruption of visual processing due to an attentional shift to the less-expected subdominant chord, response times in the visual shape identification task were expected to be slower in the less-related condition than in the baseline condition. Faster response times in the related condition than in the baseline condition would indicate that the musically related chord facilitates the simultaneous shape identification.

5. Experiment 4

5.1. Method

5.1.1. Participants

Thirty-two students of the National University of Singapore participated in Experiment 4. Number of years of instrumental instruction ranged from 0 to 16, with a mean of 3.1 (SD = 4.3) and a median of 1.0.

5.1.2. Material and procedure

Auditory stimuli were chord sequences from Tillmann et al. (2003), played with the same timbre and timing as described in Experiment 1: twelve related sequences (ending on the tonic), 12 less-related sequences (ending on the subdominant) and 24 matched baseline sequences. For the baseline sequences, the first six chords were chosen in such a way that they did not establish a tonal center: no chord was repeated and neighboring chords were harmonically unrelated (i.e., harmonically distant by at least three steps on the cycle of fifths[3]). For a given triplet of sequences (related, less-related, baseline), the last two chords were kept constant to control for local sensory influences. A baseline sequence was matched to either a related or a less-related sequence (in terms of chord types, their positions and inversions), resulting in a total of 24 baseline sequences. The baseline sequences were divided into two sets of 12 (6 matched to related sequences and 6 to less-related sequences) to present sequences of the three conditions (related/less-related/baseline) with the same frequency of occurrence (i.e., one third for each) over the experimental session. A set was presented to half of the participants.

Visual stimuli and procedure were as described in Experiment 3.

5.2. Results

Accuracy was high overall (97%, Table 2). Correct response times (Fig. 2d) were analyzed by ANOVAs with Musical relatedness (related/less-related/baseline) and Target pair (Pair A/Pair B) as within-participant factors. Only the main effects of Musical relatedness ($F_1(2,62) = 6.32$, $p < .01$, $MSE = 541.00$; $F_2(2,22) = 3.74$, $p < .05$) and of Target pair ($F_1(1,31) = 4.47$, $p < .05$, $MSE = 1109.00$) were significant. When presented together with a related chord, target shapes were identified faster than when presented with a less-related chord, as in Experiment 3, ($F_1(1,31) = 9.09$, $p < .01$, $MSE = 517.90$, $F_2(1,11) = 3.75$, $p = .08$), and also faster than when presented with a baseline chord ($F_1(1,31) = 10.00$, $p < .01$, $MSE = 551.60$, $F_2(1,11) = 8.67$, $p < .05$). No significant difference was observed between less-related and baseline conditions ($F_1(1,31) = .05$).

[3] The cycle of fifths is a music theoretical representation of keys (and their tonic chords) that is conceived of spatially as a circle. The number of steps separating two keys on this circle (whatever the direction of rotation) defines their harmonic distance.

Table 2
Percentages of correct responses for Experiment 4 presented as a function of Target pair and Harmonic relatedness (related, less-related and baseline)

Target	Harmonic relatedness		
	Related	Less related	Baseline
Pair A	95.6 (1.0)	97.7 (1.0)	97.7 (1.0)
Pair B	96.9 (1.0)	96.4 (1.0)	96.4 (1.0)

Within-participant standard errors of the mean (Masson & Loftus, 2003) are indicated in parenthesis.

5.3. Discussion

Experiment 4 replicated with a different set of chord sequences that visual shape identification is faster when a related tonic chord is presented simultaneously in comparison to when a less-related subdominant chord is presented. The comparison to the baseline condition allowed us to show that this processing difference is not caused by disruptive effects due to the less-expected subdominant chord, but rather involves facilitative effects thanks to the expected tonic chord.

6. General discussion

Previous studies have shown that musical structures influence the processing of sung syllables and words (Bigand et al., 2001; Poulin-Charronnat et al., 2005). Both language-specific interaction and general attentional processes have been proposed to explain the musical relatedness effect. Our study investigated these two hypotheses by using visual events as to-be-processed targets. The tonal function of a simultaneously played chord influenced not only the processing of visually displayed syllables (Experiments 1 and 2), but also of geometric shapes (Experiments 3 and 4). Syllable and form identification was faster when the chord played in the background was musically related than when it was less-related. Even if both mechanisms (i.e., language-specific interaction *versus* attention) are at play, only an attentional hypothesis can account for the overall data pattern.

The outcome of Experiments 1 and 2 could be explained by an interaction between language-related and musical structure processing. Together with the data for sung syllables and words (Bigand et al., 2001; Poulin-Charronnat et al., 2005), these findings support the hypothesis of shared resources for language and music processing, notably for verbal information that is either embedded in the musical information (i.e., sung) or associated with it in visual form. Numerous studies have shown shared neural resources for language and music processing, particularly for syntax (see Patel, 2003). Koelsch, Gunter, Wittfoth, and Sammler (2005) reported interactive effects for electrophysiological markers of syntax processing: irregular musical structures modified the left anterior negativity evoked by syntax errors in sentences, when these were displayed visually in synchrony to the music.

However, Experiments 3 and 4 extended the influence of musical relatedness from the processing of visually displayed syllables to geometric forms. The outcome of Experiment 3 shows that the interaction of Experiments 1 and 2 is not specific to language-related processes, but that the musical relatedness effect is mediated by attentional processes. Experiment 4 suggests that this mediation is unlikely to be due to auditory deviance detection mechanisms shifting attention away from the visual target (because of the less-related subdominant chord): the reaction times did not differ between the baseline condition and the less-related condition (i.e., with the potential deviant).

The facilitated processing with a simultaneously presented tonic chord might be due to two possible mechanisms linked to the chord's tonal function and the modulation of attentional resources available for the visual task. (1) The tonic chord is facilitated in processing because it is the most expected event in the currently established tonal center. Notably, the tonic has the highest position in tonal hierarchy (i.e., fulfilling the most important tonal function) and has the highest tonal stability. In addition, the tonic commonly serves as the final event in a musical phrase, eliciting a feeling of closure and completion, thus increasing its expectation for this position. The facilitated processing of an expected musical event would use fewer attentional resources and thus, in turn, leaving more attentional resources available for the concurrent visual processing (leading to faster visual shape identification). (2) The second possible mechanism is linked to the dynamic theory of attention (Jones, 1987; Jones & Boltz, 1989) and predicts that available attentional resources are increased for the related chord, as previously proposed for sung sequences (Bigand et al., 2001; Poulin-Charronnat et al., 2005). In this framework, auditory attention is not evenly distributed over time, but follows rhythmic attentional cycles. Melodic, tonal-harmonic and rhythmic accents work together to create the joint-accent structure of a musical sequence. This accent structure is guiding attention over time (allowing for prospective, future-oriented listening), and locally, "an accent is anything that is relatively attention-getting in a time pattern" (Jones, 1987, p. 623). Tonal-harmonic accents emerge from the established tonal center and "come from the confirmation of a contextually governed expectation about tonal relationships" (Jones, 1987, p. 623). A tonic at a phrase ending, for example, represents a "tonal end accent" because of its high tonal stability (i.e., most important tonal function) and because it marks the ending or the boundary of higher order time intervals or phrases. Because of the shortness of our sequences (i.e., eight chords with an overall duration of about 5 s), the sequences might be perceived as one overall cycle being 8-chords long, aiming for the tonic at the end (the tonal end accent), which co-occurs with some longer duration (because of the resonance of the piano timbre), thus creating a joint accent. This cycle-length seems plausible since eight beats represent the basic cycle in Jones' examples of the theoretical framework (e.g., Jones & Boltz, 1989).[4] The observed processing advantage for the tonic might thus be due to the

[4] In addition to the 8-chord cycle, shorter (4-chord) cycles might emerge in our material, at least partly. The positioning of tonal accents was not systematically controlled, but in about half of the sequences (50% in the material of Experiments 1 to 3 and 58% of Experiment 4), the fourth chord was one of the three chords at the top of the tonal hierarchy. These chords "are the mainstay of the tonality" (Piston, 1978, p. 54), also called the "harmonic core" (Bharucha & Krumhansl, 1983), and at the fourth position in the sequences, they might mark the ending of a first subgroup of four chords.

tonal end accent. The increased attentional resources in the related condition benefit to the processing of sung syllables and words (Bigand et al., 2001; Poulin-Charronnat et al., 2005) and of visual events (syllable/shape). Future research needs to further investigate this dynamic attention hypothesis, by either systematically manipulating the patterning of tonal accents in longer musical sequences (i.e., reinforcing the grouping pattern or creating different ones) or using artificially structured tone sequences without reference to the musical system.

The influence of attention modulated by the chords' tonal functions on visual performance suggests somewhat shared attentional resources between audition and vision, as also pointed out by behavioral and neurophysiological findings. In a cross-modal paradigm, a to-be-processed tone induced an attentional blink for a subsequent visual target, suggesting that attentional demands of visual and auditory tasks tap into shared resources (Jolicoeur, 1999). In a luminance detection task, target detection was facilitated when a noise burst preceded at the same location (McDonald, Teder-Salejarvi, & Hillyard, 2000). The authors concluded that visual attention is modulated by spatially congruent sounds, which facilitate low-level visual processing. Turatto, Mazza, and Umilta (2005) proposed that when auditory and visual stimuli are bound with synchronous presentations into multi-modal objects, cuing the auditory dimension of the object benefits to the processing of the non-cued visual dimension by increasing the allocated attention. The influence of cross-modal, object-based attention has been reported also for attention spreading from vision to audition (Busse, Roberts, Crist, Weissman, & Woldorff, 2005). A comparable object-based attention spread over modalities is likely to occur in our cross-modal paradigm because the conditions of synchrony and somewhat overlapping spatial locations are met (Frassinetti, Bolognini, & Ladavas, 2002; Turatto et al., 2005): chords and visual items were presented in synchrony, and at the center of the stereophonic field and of the screen. Modulation of attentional resources by the auditory information (the ending chord) affects attentional resources required for the processing of the visual information (syllable/shape).

Our study revealed an audiovisual interaction between the processing of tonal functions and visual information, which is not limited to language. The interaction observed in Experiments 1–4 can be linked to attentional resources modulated by tonal functions and possibly shared between auditory and visual modalities. This outcome has implications for the investigation of neural resources in music and language processing. Even if the data of the shape condition cannot question the existence of neural resources shared between language and music, the comparison between the data for verbal and non-verbal material suggests that at least one part of the interactions between music and language processing (syllables in Bigand et al., 2001; semantics in Poulin-Charronnat et al., 2005; syntax in Koelsch et al., 2005) is not specific to music and language, but reflects attentional processes modulated by the events' tonal functions. Control conditions using the auditory oddball paradigm (i.e., a sequence of standard tones including infrequent deviant tones) show that interactive effects are not due to perceptual deviance detection orienting listeners' attention from the visual to the auditory stimulus (Koelsch et al., 2005). However, oddball sequences do not allow evaluating the amount of interaction that is due to attentional resources modulated by tonal functions (whether it is linked to the strength of expectation or to dynamic attention guided by musical structures). In future research, control conditions using musical sequences with simultaneously presented non-verbal visual material might be useful to evaluate the specificity of the music-language-related interaction.

Acknowledgments

This research was supported in part by the grant "Junior Research Team" of the French Ministry of Research. Thanks to Bénédicte Poulin-Charronnat for comments on a previous version of this manuscript and for discussions on the hypothesis of dynamic attention.

References

Besson, M., Faïta, F., Peretz, I., Bonnel, A. M., & Requin, J. (1998). Singing in the brain: Independence of lyrics and tunes. *Psychological Science, 9*(6), 494–498.

Bharucha, J. J., & Krumhansl, C. L. (1983). The representation of harmonic structure in music: Hierarchies of stability as a function of context. *Cognition, 13*(1), 63–102.

Bigand, E., & Pineau, M. (1997). Global context effects on musical expectancy. *Perception and Psychophysics, 59*(7), 1098–1107.

Bigand, E., & Poulin-Charronnat, B. (2006). Are we experienced listeners? A review of the musical capacities that do not depend on formal musical training. *Cognition, 100*(1), 100–130.

Bigand, E., Tillmann, B., Poulin, B., D'Adamo, D. A., & Madurell, F. (2001). The effect of harmonic context on phoneme monitoring in vocal music. *Cognition, 81*(1), B11–B20.

Bigand, E., Poulin, B., Tillmann, B., & D'Adamo, D. (2003). Cognitive versus sensory components in harmonic priming effects. *Journal of Experimental Psychology: Human Perception and Performance, 29*(1), 159–171.

Boltz, M. (1989). Perceiving the end: Effects of tonal relationships on melodic completion. *Journal of Experimental Psychology: Human Perception and Performance, 15*, 749–761.

Bonnel, A. M., Faïta, F., Peretz, I., & Besson, M. (2001). Divided attention between lyrics and tunes of operatic songs: evidence for independent processing. *Perception and Psychophysics, 63*(7), 1201–1213.

Busse, L., Roberts, K. C., Crist, R. E., Weissman, D. H., & Woldorff, M. G. (2005). The spread of attention across modalities and space in a multisensory object. *Proceedings of the National Academy of Sciences of the United States of America, 102*(51), 18751–18756.

Cohen, J., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavior Research Methods, Instruments and Computers, 25*, 257–271.

Cutler, A. (1976). Phoneme-monitoring as a function of preceding intonation contour. *Perception and Psychophysics, 20*, 55–60.

Frassinetti, F., Bolognini, N., & Ladavas, E. (2002). Enhancement of visual perception by crossmodal visuo-auditory interaction. *Experimental Brain Research, 147*(3), 332–343.

Jolicoeur, P. (1999). Dual-task interference and visual encoding. *Journal of Experimental Psychology: Human Perception and Performance, 25*, 596–616.

Jones, M. R. (1987). Dynamic pattern structure in music: Recent theory and research. *Perception and Psychophysics, 41*(6), 621–634.

Jones, M. R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Review, 96*(3), 459–491.

Justus, T. C., & Bharucha, J. J. (2001). Modularity in musical processing: the automaticity of harmonic priming. *Journal of Experimental Psychology: Human Perception and Performance, 27*(4), 1000–1011.

Koelsch, S., Gunter, T. C., Wittfoth, M., & Sammler, D. (2005). Interaction between syntax processing in language and in music: An ERP study. *Journal of Cognitive Neuroscience, 17*(10), 1565–1577.

Masson, M. E. J., & Loftus, G. R. (2003). Using confidence intervals for graphically based data interpretation. *Canadian Journal of Experimental Psychology, 57*, 203–220.

McDonald, J., Teder-Salejarvi, W., & Hillyard, S. (2000). Involuntary orienting to sound improves visual perception. *Nature, 407*, 906–907.

McNamara, T. P. (2005). *Semantic Priming*. New York, NY, & Hove, UK: Psychology Press.

Neely, J. H. (1976). Semantic priming and retrieval from lexical memory: Evidence for facilitatory and inhibitory processes. *Memory and Cognition, 4*, 648–654.

Patel, A. D. (2003). Language, music, syntax and the brain. *Nature Neuroscience, 6*(7), 674–681.

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*(6), 717–733.

Peretz, I., & Coltheart, M. (2003). Modularity of music processing. *Nature Neuroscience, 6*(7), 688–691.

Piston, W. (1978). *Harmony* (4th ed.). New York: Norton.

Poulin-Charronnat, B., Bigand, E., Madurell, F., & Peereman, R. (2005). Musical structure modulates semantic priming in vocal music. *Cognition, 94*(3), B67–B78.

Port, R. F. (2003). Meter and speech. *Journal of Phonetics, 31*, 599–611.

Quené, H., & Port, R. F. (2005). Effects of timing regularity and metrical expectancy on spoken-word perception. *Phonetica, 62*(1), 1–13.

Samson, S., & Zatorre, R. J. (1991). Recognition memory for text and melody of songs after unilateral temporal lobe lesion: evidence for dual encoding. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 793–804.

Tiitinen, H., May, P., Reinikainen, K., & Näätänen, R. (1994). Attentive novelty detection in humans is governed by pre-attentive sensory memory. *Nature, 372*, 90–92.

Tillmann, B., Bharucha, J. J., & Bigand, E. (2000). Implicit learning of tonality: a self-organizing approach. *Psychological Review, 107*(4), 885–913.

Tillmann, B., Janata, P., Birk, J., & Bharucha, J. J. (2003). The costs and benefits of tonal centers for chord processing. *Journal of Experimental Psychology: Human Perception and Performance, 29*(2), 470–482.

Turatto, M., Mazza, V., & Umilta, C. (2005). Crossmodal object-based attention: Auditory objects affect visual processing. *Cognition, 96*(2), B55–B64.