Processing pitch and duration in music reading: a RT–ERP study

Daniele Schön [a,b,*], Mireille Besson [a]

[a] *Equipe Langage et Musique, Centre de Recherche en Neurosciences Cognitives (CRNC), CNRS*
31, Chemin Joseph-Aiguier, 13402 Marseille Cedex 20, France
[b] *Dipartimento di Psicologia, Università di Trieste, Trieste, Italy*

Abstract

The originality of this study is to examine the processing of pitch and duration in music reading, using both electrophysiological and behavioral methods. Specifically, it was of interest to determine whether pitch and duration in music reading are processed independently or jointly. A probe, comprising a key and time signature was presented, and participants were required to compute the tonic and/or the best fitting duration. A target note followed the probe and participants made a match/mismatch judgment on the dimension they were required to analyze (i.e. pitch or duration). We hypothesized that, if pitch and duration are processed independently results should show no interference of the irrelevant dimension on the relevant. Results showed that early differences emerged in the ERPs as a function of the task to be performed on the target in block 1. Moreover, RTs were shorter in the pitch than in the duration task and for congruous than incongruous targets. In the ERPs, this congruity effect was reflected by a negative component, to incongruous targets. Most importantly, the congruency of the target note on the irrelevant dimension did not have any effect on the ERPs, suggesting that pitch and duration are processed independently. © 2002 Elsevier Science Ltd. All rights reserved.

Keywords: Music notation; Temporal; Tonal; Metric; Pitch

1. Introduction

There is a general agreement in identifying pitch and temporal structures as two main dimensions in music. From a musical point of view one could argue that, since melody and harmony both contribute to the rhythmic organization of a musical work, and since neither melody nor harmony can be activated without rhythm, the three must be regarded as inseparably linked. That is to say, a complete representation of music requires an integration of these dimensions. Certainly, the weight of every single dimension on the final representation will depend upon musical culture, period or style considered. For example, much of western music theory, from counterpoint to twelve-tone harmony, is concerned with the pitch dimension [36]. This might explain the frequently repeated observation that time in music has received rather less attention than pitch [9]. Still it should be kept in mind that temporal patterns are of major importance in some musical cultures [4].

The literature on music perception in the auditory domain highlights two different positions regarding the relationship between pitch and temporal structures. Some argue that pitch and time are jointly encoded into the cognitive system, whereas others claim they are encoded independently. Interactions between pitch and temporal processing have been demonstrated many times in cognitive psychology research. Pitch changes were found to be easier to detect in a context of predictable temporal structure [21]. Melodies were judged to unfold more slowly when they displayed more changes in pitch direction and greater pitch interval [5]. Boltz [6] also found an interaction effect, even when she systematically manipulated participants' attention to the temporal and pitch properties of melodies. Moreover, memory recall performance diminishes when pitch and temporal structures are not in phase [7,10,18]. Recently, neuroimaging techniques have demonstrated interactions between the two dimensions [14,27].

In contrast, following the second position, the neural mechanisms underlying pitch and rhythm processing would be independent. Strong evidence supporting this hypothesis comes from neuropsychological studies. Peretz [29] reported that two right hemisphere damaged patients showed a selective deficit in the processing of pitch, while a reverse pattern of selective deficit of the temporal dimension was found in two left brain damaged patients. Other single case studies have also demonstrated selective impairment of one

* Corresponding author. Tel.: +33-4-91-16-41-13;
fax: +33-4-91-77-49-69.
E-mail address: danschon@lnf.cnrs-mrs.fr (D. Schön).

or the other dimension [13,24,30,31], and other studies in the field of cognitive psychology support the hypothesis of independence [15,26,28].

1.1. Music reading

While many studies have been devoted to understanding the relationship between pitch and temporal processing in the auditory modality, the originality of this research is to investigate this relation in the visual modality. Music can be approximately represented in notational form.[1] History of western notation begins in the ninth century and it still continues to evolve [34]. Pitch and temporal information are represented in western modern notation differently one from the other. The diasthematics (from Greek diasthema: interval) of staff notation, that is the pitch dimension, is in part ordinal and in part intervallic [22]. It is ordinal in the way higher notes are represented proportionally higher on the staff. At the same time, it can be considered intervallic since different graphical distances on the vertical dimension reflect proportional musical intervals. The form of the note represents time and can vary according to three parameters that may be combined in different ways: the stem, the head of the note and the dot(s) after the note. The temporal dimension unfolds horizontally and it is intervallic to the extent that the space between one note and the following one is proportional to the duration of the note. Therefore, at least in music notation, pitch and temporal information are coded separately, distinguished by different attributes.

Findings in the neuropsychological literature also support evidence for dissociation between pitch and temporal processing in music reading, in analogy with those reported in the auditory domain. Assal [1] reported the case of a patient with Wernicke's aphasia and mild music dyslexia, who had much more difficulty in rhythm reading than in pitch. In contrast, Brust [8] described a patient with a selective disturbance of pitch reading, but with preserved rhythmical competence. Fasanaro et al. [12] described the case of a patient who had a musical alexia and agraphia limited to the pitch component, with a preserved processing of rhythmic structures. They argue that the pitches in a music score are analyzed separately from rhythmic and ideographic symbols. Horikoshi et al. [16] report that the recognition of the pitch and rhythm of notes in music reading were dissociated in a patient who showed a selective deficit in naming or playing at the piano even simple four notes sequences.

Within the psychological literature, most of the studies in music reading are concerned with the ability of sight-reading (i.e. playing for the first time a music score), but little research has been aimed at investigating the relationship between pitch and temporal processing. In a recent study conducted by Waters et al. [37], pitch and temporal dimensions were modulated in a matching task by varying the coherency of the two dimensions (coherent versus random). They found that the effects of pitch and temporal structures on expert performance were quite weak, probably due to the small difference between coherent and randomized sequences. In another study, the same researchers [38] using a visual stimuli recall paradigm, found that rhythmic auditory distractors did not disrupt the written recall of temporal information, while rhythmic and pitch auditory distractors had the same disrupting effect on pitch recall. They thus conclude that, at least in the written recall task they used, pitch processing is constrained by the processing of temporal structures, but not vice versa. This conclusion is not supporting either strong interactive or independence models of pitch and time processing.

In this work, we focused on the early stages of pitch and temporal information processing in music reading. Participants were presented a signature[2] as a probe that comprised both tonal and metrical information. They were required to decide, as quickly and as accurately as possible, whether the single target note that followed probe presentation, matched or not with the tonal or metrical information provided by the probe. Three main factors were manipulated: task (pitch or duration), congruity of the relevant dimension (congruous or incongruous), congruity of the irrelevant dimension (congruous or incongruous).

Our specific aim was to determine whether pitch and duration are processed as two separate dimensions or whether they are, on the contrary, processed as integrated dimensions. By presenting the key and time signature on the staff (probe), we hoped to induce strong expectancies regarding the pitch and duration dimensions of the target note. As participants were required to judge only one dimension (pitch or duration) of the target note, it was, therefore, possible to determine whether or not the irrelevant dimension produced an interference effect on the relevant dimension. If pitch and duration are processed independently, results should show no interference (e.g. in the pitch task, reaction times, RTs, to congruous targets on the relevant dimension, pitch, should not differ significantly as a function of whether the target is congruous or incongruous on the irrelevant dimension, duration). If this interference were to be found, its effect should be reflected in longer reaction times, as demonstrated many times in Stroop-like paradigm [23].

In the first part of the experiment (block 1), subjects were required to focus attention only on one dimension of the probe. Results therefore allow making inferences on the pro-

---

[1] We use the term approximately because, although expressive indications may appear on a musical score, there are many aspects of music that cannot be precisely represented in notational form (e.g. the notes of a chord in a piano score may be played simultaneously or successively depending on the context and on the interpretation).

[2] In this case the violin (G) clef was placed on the traditional five-lined staff, with accidentals, indicating the precise tonality, on the right (in a broad sense tonality denotes a system of relationships between pitches having a 'tonic' or central pitch as its most important element). Indications regarding the temporal structure were presented on the right of accidentals, taking the form of a numerical fraction: the denominator indicates the chosen unit and the numerator the number of units that constitute the measure.

cessing of tonal and metric information of the probe. How-
ever, finding of no effect of the irrelevant dimension of the
target could be explained by the fact that the subjects, hav-
ing completely ignored the irrelevant feature of probe, did
not have any expectation on that dimension of the target.
On the contrary, in block 2 subjects were forced to pro-
cess both dimension of the probe, in a way they were re-
quired to build a representation of both dimensions of the
note.

To study the relationship between pitch and duration
processing, we used both RT measures and event-related
potentials (ERPs) methodology. To our knowledge, no ERPs
study of the processes involved in music reading has yet
been conducted. Three main reasons led us to choose the
ERP method, together with RTs, to address the questions
above. First, the ERP method is known for its excellent
temporal resolution: it allows determining, for instance,
whether two processes are qualitatively similar or different
and if different, when they start to diverge. Second, it was
of interest to study the congruity effect on the relevant di-
mension (that is, the difference between incongruous and
congruous target notes). It is well known, from previous
results in language experiments, that a negative component,
peaking around 400 ms, is elicited by incongruous words
relative to a linguistic context [19]. As far as music is con-
cerned, incongruous or wrong notes at the end of musical
phrases are generally associated with positive components,
peaking around 600 ms [2]. It was therefore of interest
to determine whether target notes incongruous relative to
the key and time signature (probe), would elicit an N400
component, a P600 component or yet another effect, and if
this effect would be similar or different for the pitch and
duration dimensions. Finally, the ERP method should also
provide interesting and complementary information regard-
ing the independent or integrated processing of pitch and
duration. If pitch and duration are processed independently,
the congruity effect on the relevant dimension should not
be influenced by the congruity effect on the irrelevant
dimension.

2. Method

2.1. Participants

Eighteen volunteer musicians (mean musical expertise of
16 years) were tested individually in an experiment that
lasted for approximately 2 h. All were right-handed, neuro-
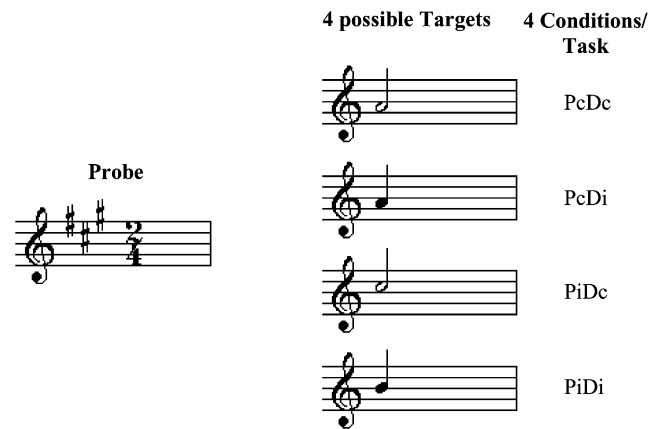logically normal, had normal or corrected to normal vision,

Fig. 2. Illustration of the experimental conditions: Pc, pitch congruous;
Pi, pitch incongruous; Dc, duration congruous; Di, duration incongruous.

and were native French speakers (mean age = 25.5 years).
None of them had absolute pitch. All participants were paid
for their participation to the experiment. One participant per-
formed poorly and was thus excluded from further analysis.
Moreover, two other participants were also excluded from
further analyses due to the large number of electrophysio-
logical artifacts.

2.2. Stimuli

Stimuli were musical symbols presented visually at the
center of a computer screen (SVGA color computer screen
placed 60 cm in front of the participants). A staff with key
and time signature (i.e. the violin clef and the alterations:
flats or sharps, signs defining the tonality) and metric in-
dications specifying the unit and number of beats/measure
(e.g. 2/4, 3/8, 3/4, etc.) was presented for 500 ms (see
Fig. 1). The screen was blank for 1000 ms and a target
note was then presented on the staff for 500 ms. A row of
X was presented 1500 ms after target offset for 2000 ms
(inter-trial-interval). Four experimental conditions were
used (see Fig. 2): (1) the target note was congruous with the
indications regarding both tonality and meter (PcDc); (2)
the target note was congruous with the indications regard-
ing tonality, but not with meter (PcDi); (3) the target note
was congruous with the indications regarding meter, but not
tonality (PiDc); (4) the target note was not congruous either
with tonality or meter (PiDi). There were 60 trials in each
experimental condition. The 60 probe and 60 target stimuli
were used in blocks 1 and 2 (6 pitches × 5 duration × 2
colors).

Fig. 1. Experimental design.

2.3. Procedure

Participants, comfortably seated in a Faraday box, were instructed to decide, as quickly and as accurately as possible, whether the target note matched or mismatched with the tonal or metrical information given by the probe.

In block 1, either the accidentals or the numbers of the probe were colored (green or red) and participants were required to judge whether the target note matched or not the colored dimension of the probe. Thus, they knew in advance which task, pitch or duration, they would have to perform upon target presentation. In the pitch task, participants were asked if the target note was the most representative of the key/tonality previously established by the probe (e.g. a D matching two sharps, only major tonalities being taken into account). In the duration task, they judged whether the note properly fitted the duration of the measure, as indicated by the metric information of the probe (e.g. a half note matching 2/4).

Block 2 was similar to block 1, but the probe was not colored and therefore did not provide any information on the task to be performed. The target note was presented in red or green and provided this information. Depending upon the color of the target, musicians were asked to perform the pitch task or the duration task.

All participants performed the two blocks successively and in the same order.[3] Within participant the same color/task association was used in both blocks. This color-task association was counter balanced across participants (with accidentals colored in red for half the participants and in green for the other half and vice versa for the numbers). Also the association between fingers (index, middle) and responses (yes or no) was balanced across participants.

Each block began with a practice session to familiarize participants with the tasks and to train them to blink during the ITI, while the XXX appeared on the screen. During the ITI an acoustic feedback, signaling incorrect answers, was given as well. A total of 240 trials were presented in each block, with 120 trials in the pitch and 120 trials in the duration task. Pitch and duration tasks were intermixed within a block using a pseudo-random presentation (not more than three repetitions of the same task).

2.4. Design

These factors were manipulated within participants: the task requirements (pitch versus duration judgments), the congruity on the relevant dimension (congruous/incon-gruous) and the congruity of the irrelevant dimension (congruous/incongruous).

2.5. Data acquisition and analysis

EEG was recorded for 3700 starting 200 ms before the onset of the probe (baseline), from 28 scalp electrodes, mounted on an elastic cap, and located at standard left and right hemisphere positions over frontal, central, parietal, occipital, and temporal areas (international 10/20 system sites: Fz, Cz, Pz, Oz, Fp1, Fp2, F3, F4, C3, C4, P3, P4, O1, O2, F7, F8, T3, T4, T5, T6, Fc5, Fc1, Fc2, Fc6, Cp5, Cp1, Cp2, Cp6). These recording sites plus an electrode placed on the right-mastoid were referenced to the left-mastoid electrode; the data were then re-referenced offline to the algebraic average of the left- and right-mastoids. Impedances of these electrodes never exceeded $3 \, k\Omega$. The horizontal electrooculogram (EOG) was recorded from electrodes placed 1 cm to the left and right of the external canthi, and the vertical EOG was recorded from an electrode beneath the right eye, referenced to the left-mastoid, to detect blinks and vertical eye movements. Trials containing ocular artifacts, movement artifacts, or amplifier saturation were excluded from the averaged ERP waveforms. The mean rejection rate was 21%, with a maximum of 35% of rejected trials for a given participant. The EEG and EOG were amplified by an SA instrumentation amplifier with a band pass of 0.01–30 Hz, and were digitized at 250 Hz by a PC-compatible microcomputer (Compaq Prosignia 486).

ERP data were analyzed by computing the mean amplitude in selected latency windows relative to a 200 ms pre-probe-display baseline. Analysis of variance (ANOVA) was used for all statistical tests, and all *P*-values reported below were adjusted with the Greenhouse–Geisser epsilon correction for non-sphericity. Reported are the uncorrected degrees of freedom, the epsilon value, and the probability level after correction. Unless specified, *t*-tests were used to test the significance of post hoc comparisons. Only trials in which performance was correct were taken into account in the analyses. Two types of ANOVAs were performed, that included either all the electrodes or the midline electrodes only. Results were very similar in both analyses. Moreover, the ANOVA including all electrodes did not reveal any strong localization pattern of the effects of interest on the scalp. Therefore for simplification purposes, we only report the results for the midline electrodes.

3. Behavioral results

All participants were able to perform the two blocks, though block 1 was definitely easier than block 2 (mean percentage of error in block 1 = 3.22, S.D. = 2.2, mean percentage of error in block 2 = 12.6, S.D. = 7.3; *t*-test $P < 0.0001$).

---

[3] Two main reasons led us to do so. First and most importantly, in block 1 we wanted subjects to focus attention on one dimension only of the probe. If participants performed the block 2 first (both dimensions must be processed), it would have been difficult to ensure that they would not use the same strategy in block 1. Second, as the task in block 2 is somewhat more difficult to perform, block 1 provided some training.
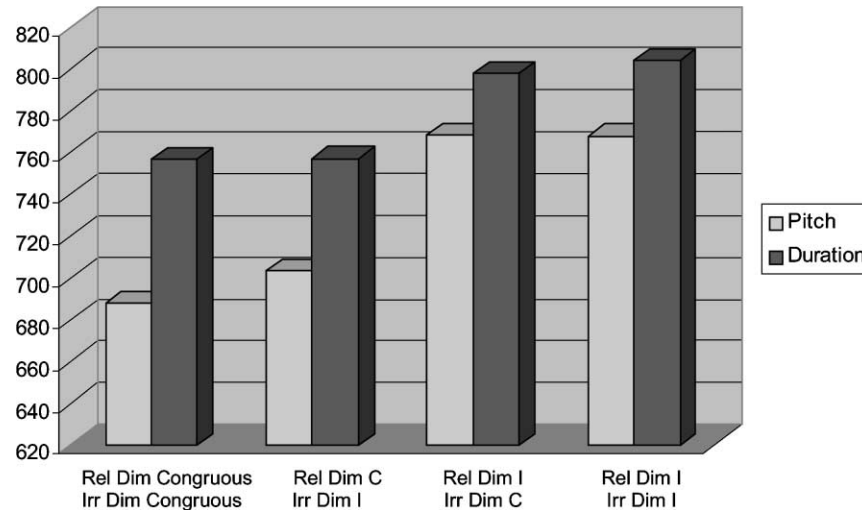
Fig. 3. Mean RT for each condition in each task in block 1. RTs (ms) are reported for the four experimental conditions for block 1 (left to right in abscissa): both dimensions congruous; relevant dimension congruous and irrelevant incongruous; relevant dimension incongruous and irrelevant congruous; both dimensions incongruous. RTs are plotted separately for the pitch and duration tasks.

3.1. Block 1

A three-way repeated measure ANOVA was carried out including task (pitch/duration), congruity on the relevant dimension (congruous/incongruous) and congruity on the irrelevant dimension, as factors. The main effect of task was significant ($F(1, 14) = 12.08$, MSE $= 66,259$, $P = 0.004$), with shorter RTs (Fig. 3) in the pitch task (731 ms) than in the duration task (778 ms). RTs to targets congruous on the relevant dimension (726 ms) were shorter than to incongruous (784 ms) targets ($F(1, 14) = 19.44$, MSE $= 101,660$, $P = 0.001$). Moreover, the task × congruity on the relevant dimension interaction was significant ($F(1, 14) = 5.14$,

MSE $= 6216$, $P = 0.04$): the effect of task was larger for congruous than incongruous targets. The *t*-tests revealed that pitch was significantly faster than duration task when the relevant dimension was congruous with the probe ($P = 0.002$). However, when the relevant dimension was not congruous, significance was reached only when the irrelevant dimension was also incongruous ($P = 0.043$), but not when the irrelevant dimension matched the information carried by the probe ($P = 0.16$). Finally, the main effect of the irrelevant dimension was not significant ($F(1, 14) = 1.83$, MSE $= 906$, $P = 0.19$) and there were no interaction with the effects of the other two factors ($F < 1$, in both cases).
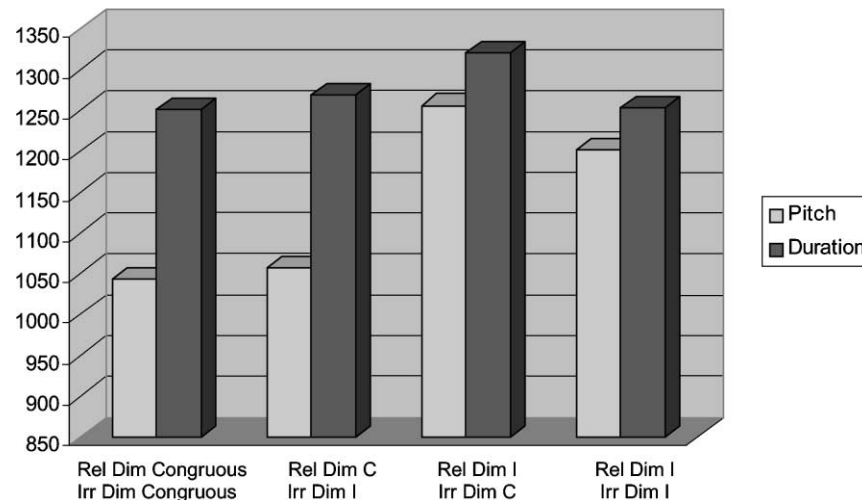

Fig. 4. Mean RT for each condition in each task in block 2. RTs (ms) are reported for the four experimental conditions for block 2 (left to right in abscissa): both dimensions congruous; relevant dimension congruous and irrelevant incongruous; relevant dimension incongruous and irrelevant congruous; both dimensions incongruous. RTs are plotted separately for the pitch and duration tasks.

3.2. Block 2

Overall results were similar in blocks 1 and 2 (see Fig. 4). Again, a three-way repeated measure ANOVA was carried out. The main effect of task was significant ($F(1, 14) = 10.30$, MSE = 489,675, $P = 0.006$), with shorter RTs in the pitch task (mean = 1139 ms) than in the duration task (mean = 1273 ms). RTs to congruous targets on the relevant dimension (1155 ms) were shorter than to incongruous targets (1257 ms; $F(1, 14) = 27.9$, MSE = 313,387, $P < 0.001$). Moreover, the task × congruity on the relevant dimension interaction, was also significant ($F(1, 14) = 8.30$, MSE = 156,521, $P = 0.012$). Again t-tests revealed that pitch was significantly faster than duration when the relevant dimension was congruous with the probe ($P = 0.005$). However, RTs in pitch and duration tasks were not different when the relevant dimension was incongruous with probe information ($P = 0.17$). Finally, while

the main effect of the irrelevant dimension was not significant ($F(1, 14) = 1.00$, MSE = 14,578, $P = 0.33$), the interaction between the effects of congruity on the relevant and irrelevant dimension was significant ($F(1, 14) = 9.30$, MSE = 48,897, $P = 0.009$). However, detailed statistical analysis (t-tests) only showed a small, non-significant interference effect of the irrelevant dimension on the relevant dimension when targets were incongruous ($P < 0.09$), and no interference effect when targets were congruous on the relevant dimension ($P > 0.5$).

4. ERPs results

4.1. Block 1

Four-way repeated measures ANOVAs including task, congruity on the relevant dimension, congruity on the irrel-
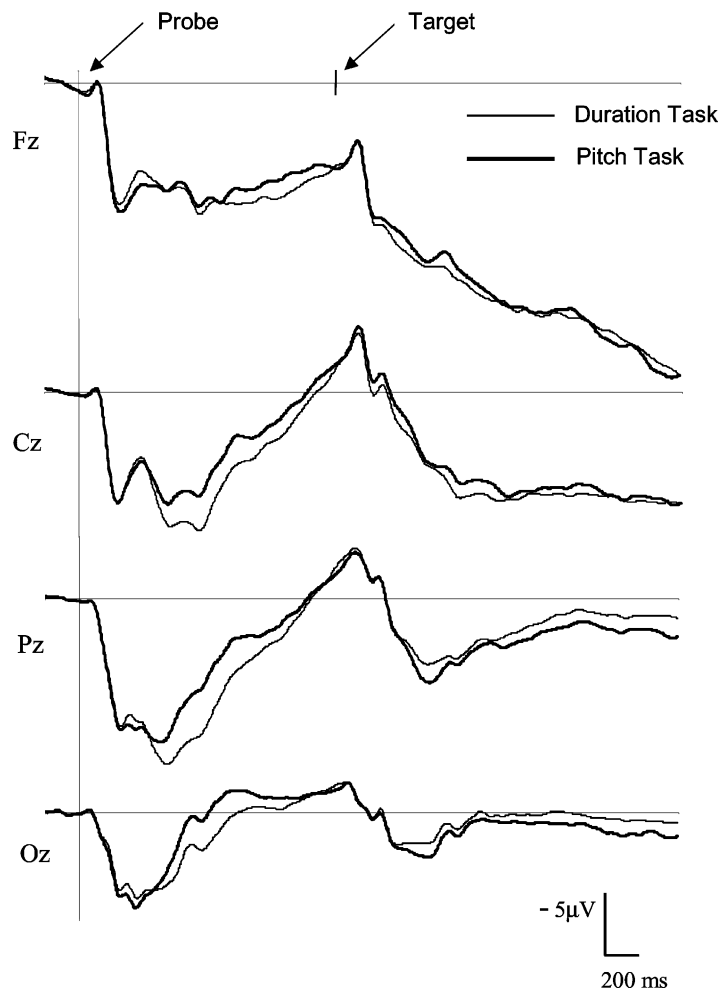

Fig. 5. Effect of task in block 1. ERPs results, averaged across 15 musicians and across the four experimental conditions, are presented separately for the duration task and for the pitch task, in block 1. Recordings are from the midline electrodes, at frontal (Fz), central (Cz), parietal (Pz) and occipital sites (Oz). ERP were recorded for 3700 ms, starting 200 ms before the onset of the probe, marked by the vertical line. During the probe-target interval, ERPs are more positive in the duration than in the pitch task (400–1300 ms). In this and subsequent figures, amplitude (μV) is represented on the ordinate, with negative voltage up, and time (ms) on the abscissa.

evant dimension, and electrodes (midline) as factors, were performed in successive latency bands. In the 0–400 ms range, results showed no significant main effect of task ($F(1, 14) = 0.18$, MSE $= 5.18$, $P = 0.67$), suggesting that the task did not influence the very early stages of processing (Fig. 5). Results in the 400–1300 ms range showed a significant main effect of task ($F(1, 14) = 15.26$, MSE $= 292.92$, $P = 0.0016$) with larger positivity in duration than in pitch processing. None of the interaction including the task factors was significant. In the 200 ms previous to the target, that is to say in the 1300–1500 ms range, the main effect of the task was no longer significant and the ERP traces very well overlapped ($F(1, 14) = 0.60$, MSE $= 16.17$, $P = 0.45$). After target onset, the effect of task was not significant in any of the latencies bands considered for analysis. However, the main effect of congruity on the relevant dimension was significant in the 1800–2000 ms range (that is, 300–500 ms after target onset; $F(1, 14) = 23.04$, MSE $= 981.49$, $P =$ 0.0003) with a negative-going component associated with incongruous compared to congruous target notes (Fig. 6). Finally, the main effect of the irrelevant dimension was not significant, nor were any of the interactions involving this factor. Note also that none of the interactions involving the effect of electrodes was significant in any of the latency bands considered for analysis.

4.2. Block 2

In the second block, participants did not know until target presentation which task they were required to perform. Thus, they were compelled to process both tonal and metrical information given by the probe. Accordingly, results showed no significant difference as a function of task before the target was presented. Results of four-way repeated measures ANOVA, including the same factors as in block 1, revealed a main effect of congruity on the relevant
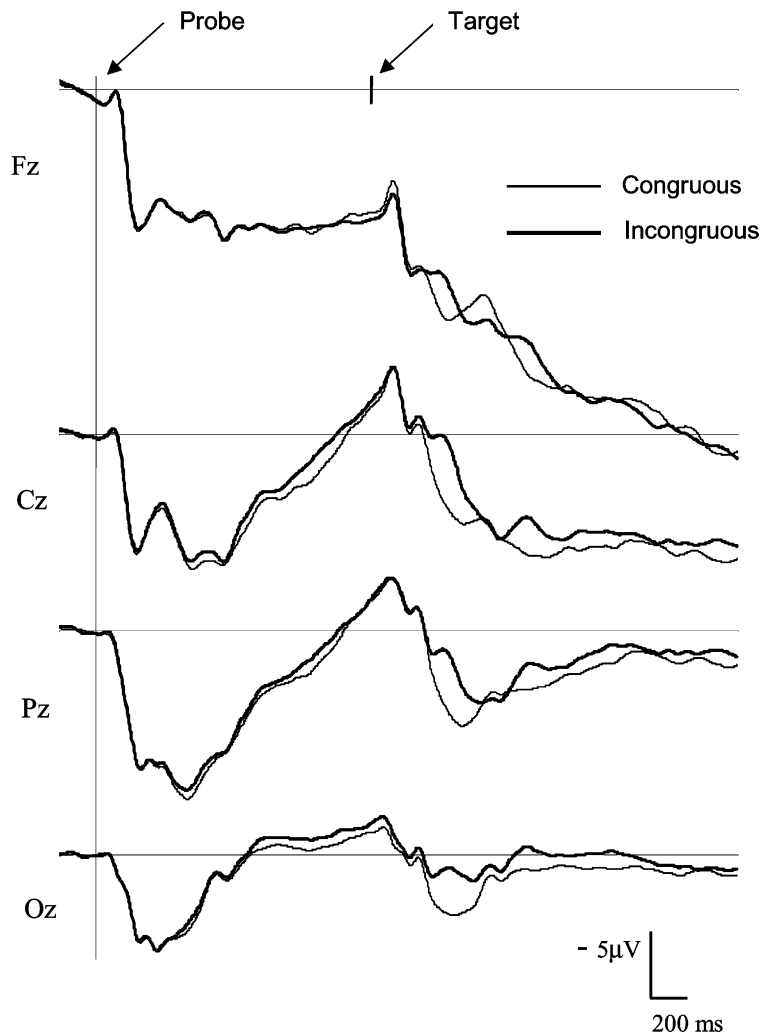

Fig. 6. Effect of congruity on the relevant dimension in block 1. ERPs results, averaged across 15 musicians and across the two tasks (pitch and duration) in block 1, are presented separately for targets that are congruous or incongruous on the relevant dimension. ERPs to incongruous targets are relatively more negative than ERPs to congruous targets, peaking 400 ms after target onset.
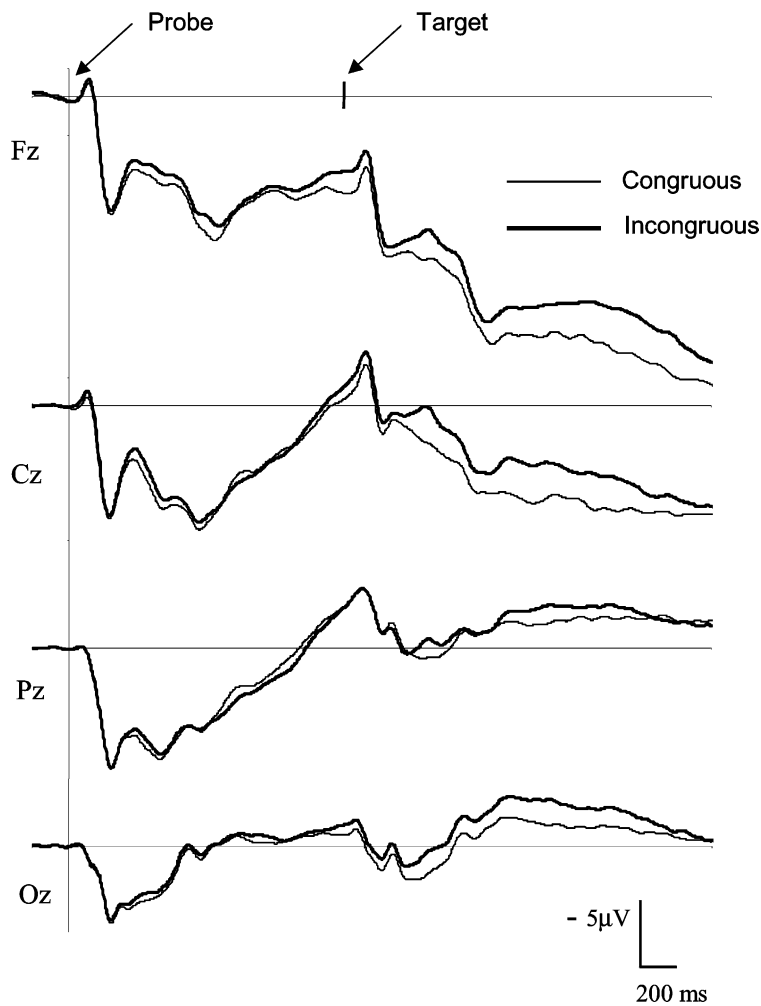
Fig. 7. Effect of congruity on the relevant dimension in block 2. ERPs results, averaged across 15 musicians and across the two tasks (pitch and duration) in block 2, are presented separately for targets that are congruous or incongruous on the relevant dimension. Again, ERPs to incongruous targets are relatively more negative than ERPs to congruous targets, peaking 400 ms after target onset.

dimension. The congruity effect (Fig. 7) was significant in the 1800–3100 ms range, that is between 300 and 1600 ms after target onset ($F(1, 14) = 27.74$, MSE = 1740, $P = 0.0005$). No congruity × task interaction was found ($F < 1$). Finally, the main effect of the irrelevant dimension was not significant, nor were any of the interactions involving this factor (F is always <1.7). Note also that none of the interactions involving the effect of electrodes was significant in any of the latency bands considered for analysis.

5. Discussion

The study was divided in two experimental blocks that differed by the amount of information provided by the probe. In block 1, the accidentals or the numbers on the staff were colored. Therefore participants knew, upon probe presentation, which dimension was relevant to perform the task. They could use this information to selectively process the relevant dimension and prepare their response on the target. In block 2, in contrast, the probe did not provide any information on the relevant dimension. To be able to decide whether the target matched or mismatched with the probe, participants therefore had to process both the pitch and duration information.

Overall, results showed that differences emerged in the ERPs depending upon the computations required to process the pitch or duration information provided by the probe in block 1. Moreover, RTs were shorter in the pitch task than in the duration task and for congruous than incongruous target notes. In the ERPs, this congruity effect was reflected by a negative component to incongruous targets, peaking around 450 ms. Most importantly, the congruency of the target note on the irrelevant dimension did not have any effect on the ERPs, neither in block 1 or in block 2, and had only a small effect on RTs in block 2.

Let us first consider results concerning probe processing in block 1. When the probe provided advance information

on the dimension (pitch or duration), relevant to perform the match/mismatch decision on the target (block 1), ERPs in the two conditions diverged 400 ms after probe onset. The finding of no significant differences from probe onset to 400 ms later seems to indicate that this reflected the time required to analyze the physical properties of the probe staff and to focus attention on the relevant dimension. Then, the ERPs diverged, in the 400–1300 ms latency range, with significantly smaller positivity in tonal processing compared to metric processing. The difference between pitch and duration may be considered either as an increased positivity in the duration task or as an increased negativity in the pitch task. We interpret our results as showing a relative increase in negativity in the pitch task, because similar results in the literature have often been considered as reflecting a greater activation of processing resources due to higher processing demands [3,17]. In fact, this interpretation fits well with music theory and the comments of the musicians upon debriefing at the end of the experiment. Indeed, in music theory, tonality computations are generally considered as more demanding than metric computations. The pitch task requires extracting the tonality from the accidentals on the staff. Such computations may be realized using two different strategies. The first strategy consists of analyzing the last (to the right) sharp and adding half a tone to get the tonic that is then compared to the target (or taking the note two-fifths above the last sharp). For tonalities with flats, the second last flat (or the note a fifth below the last) coincides with the tonic. The second strategy is more "automatic"; it consists of directly accessing the tonality from a given number of accidentals (e.g. three sharps define A as major tonality). Interestingly, these two strategies have been shown to be at play in arithmetic calculation as well [25]. Thus, in analogy with number processing, if 3×4 is considered as an arithmetic fact, three sharps can be considered as an 'harmonic fact', since in both cases the problems are solved by directly accessing the solution, without the activation of calculation procedures [25]. In the duration task, preparing for the response does not require many computations. Musicians most likely access 'metric facts' directly to associate metric information with specific notes' duration.

It is important to note that the ERP differences between tonal and metric processing were no longer significant in the 200 ms latency band preceding the target. Therefore, the RT differences found to the targets are unlikely to be related to the ERP differences seen to the probe. Rather they are related to note identification and comparison with the representation built from the probe. Three, not mutually exclusive, interpretations can account for the fact that pitch judgment was faster than duration. First, in written music, pitch changes more frequently than duration. Thus, in a series of notes, the probability that two following notes change in pitch is higher than they change in duration. This is especially true in western music in which harmonic patterns are more salient than rhythmic patterns. Second, in reading/playing a score, a musician needs to anticipate the correct pitch before the onset

of the note. In contrast, duration is determined by the offset of the note, thus it is second in temporal priority. Third, pitch identification might be more language mediated than duration. Indeed, the pitch of notes can easily be named (A, B, etc.), while naming duration is not so frequent and it may also require compound words (dotted half note, etc.). Thus it may be that symbolic language mediation facilitated pitch identification and matching judgment.

The finding that RTs were shorter for congruous than incongruous targets is in line with many results in the literature. In our experiment, a particular target note, with a specific pitch or duration, was expected on the basis of probe information. Thus, while participants could anticipate congruous targets, and therefore prepare for their response, they could not anticipate incongruous targets, which may explain why RTs were slower in this latter case. Moreover, while the ERPs to congruous and incongruous targets well overlapped from target onset to 300 ms later, that is when sensory/perceptual processing is taking place, incongruous target notes were associated with a negative-going component that developed in the 300–500 ms latency band after target onset.

In our experiment, musicians could build specific expectations based upon probe information. The processes called into play may therefore be quite similar to those involved in a semantic priming paradigm, in which, typically, participants have to decide whether a target word (e.g. apple, wind) does or does not belong to the semantic category defined by the probe (e.g. fruit). Three criteria are generally used in ERPs research to disentangle components: their latency, scalp distribution and sensitivity to experimental variables [11]. Relative to the sensitivity to experimental variables and the latency criterion, one may argue that the negative component found in the present experiment belongs to the N400 family. Indeed, N400 components are typically found when a word is unexpected within a semantic context, be it another word or a sentence [19]. However, the fronto-central scalp distribution of the negativity reported here does not correspond to the typical centro-parietal distribution of the N400 component in the visual modality [20].

In contrast, the fronto-central distribution of the negativity found here is reminiscent of the scalp distribution of the N200 component that has been widely described in the ERP literature as reflecting a general mismatch effect [32,33]. The occurrence of the N200 is often linked to a subsequent positivity, P3a, and forms the N2–P3a complex, which has been shown to develop in response to surprising, unexpected stimuli. The findings that the negativity reported here is also followed by a positive component, and that similar mismatch effects are found in both pitch and duration tasks (no task by congruity interaction), may therefore be taken as evidence for considering our results as an N2–P3 complex rather than as an N400 component. Further research is clearly needed to strengthen this interpretation.

Finally, and most importantly for the purpose of the present experiment, in block 1 analyses of the ERPs results

did not show any influence of the irrelevant dimension on the relevant dimension in any of the latency bands considered for analysis and any interaction with other factors. Nor did RT results. Thus, these results seem to favor the view that pitch and duration are processed independently. However, one could argue that insofar as musicians knew the relevant dimension from probe onset, they could focus attention on one dimension of the probe and completely ignore the other, thus avoiding any interference effects in target processing. It is therefore important to compare results in block 1 with those in block 2, in which no information regarding the relevant dimension was provided before target onset. Upon probe presentation, participants had to process both pitch and duration dimensions. Therefore, interference effects, if any, were more likely to be found in block 2 than in block 1. However, ERPs results in block 2 did not show any influence of the irrelevant dimension on the relevant dimension. Similarly, RT results showed no main effect of the irrelevant dimension and no interaction with the effects of the task factor. The only result that possibly contradicts the independence hypothesis is that congruity on the relevant dimension significantly interacted with congruity on the irrelevant dimension. Specifically, RTs to incongruous targets on the relevant dimension were somewhat faster ($P < 0.09$) when the targets were also incongruous on the irrelevant dimension than when they were congruous. One possible explanation is linked with the RT difference between congruous and incongruous targets. Insofar, as RTs were overall longer for incongruous than for congruous targets, there was more time for an interference effect to develop when the target notes were incongruous than when they were congruous. In line with this interpretation, it is interesting to note that no interference effect of the irrelevant dimension was found in block 1, in which, overall, RTs were faster than in block 2 (by 450 ms).

Taken together, results clearly favor the hypothesis that pitch and duration are processed independently, at least during the early stages of information processing that we examined in the present study. However, some caution should be taken, at least for two reasons. First, analyses of the ERP results did not indicate any topographical differences as a function of task. Therefore, one could argue that similar neural networks were at play in the two tasks. However, it may also be that the number and the configuration of electrodes chosen for the analysis were not sensitive enough to reveal such differences. To further track this issue, other brain imaging methods better suited for spatial localization, such as functional magnetic resonance imaging, should be used in further experiments. Second, one could argue that the irrelevant information was not salient enough to interfere with performance. To address this issue, a study needs to be conducted in which the saliency of the irrelevant dimension is manipulated. Then, one would expect that the magnitude of the deviancy of the task-irrelevant information would have little impact on processing the task-relevant feature. Finally, it is interesting to compare our results with

those reported by Waters and Underwood [38] in an experiment on music reading. As we earlier described, their results provide partial evidence in favor of an independent processing of pitch and duration, but they do not strongly support either an independent or an interactive model of pitch and duration processing. It is possible to reconcile our findings with theirs by taking into consideration the differences between the tasks used in the two studies. In the Waters and Underwood studies visually presented sequences of notes were encoded for a delayed recall test, while, in our experiment, a single note was encoded for an immediate match/mismatch decision task. If it is indeed the case that interference effects from the irrelevant dimension need time to develop, it is therefore more likely to find such interference effects in the delayed recall task than in the immediate match/mismatch decision task. Furthermore, when reading a score, a musician is most likely to use one of three strategies, depending on the task at hand: singing-like (i.e. visual to auditory transcoding), playing-like (visual to motor transcoding) or naming-like notes (visual to verbal transcoding). The degree of interaction between these strategies may vary according to the task and to the musician [35]. The use of a written recall task by Waters and Underwood [38] may have encouraged the development of an auditory strategy (visual to auditory transcoding) since their results showed that auditory interferences partially disrupted the recall of visual stimuli. In our study, auditory and/or motor transcoding strategies were not likely to be used, due to the use of single note as targets. Generally speaking, to be processed, single notes do not require any powerful strategy such as visual to auditory coding that may be necessary when long sequences of notes must be retained in memory. Moreover, notes acquire a real musical sound and duration only when they are coded one in relation to one another. In other words, in absence of a frame of reference, reading a single note is not likely to favor an auditory coding, except, may be for musicians with absolute pitch.

In conclusion, these results speak to the question as to whether pitch and duration can be processed independently. The participants' ability to focus attention on one dimension while neglecting the other indicates that the processes involved in each of these dimensions are, in all likelihood, structurally and functionally independent.

References

[1] Assal G. Aphasie de Wernicke chez un pianiste. Revue de Neurologie 1973;29:251–5.

[2] Besson M, Macar F. An event-related potential analysis of incongruity in music and other non-linguistic contexts. Psychophysiology 1987;24:14–25.

[3] Birbaumer N, Elbert T, Canavan AG, Rockstroh B. Slow potentials of the cerebral cortex and behavior. Physiological Review 1990;70:1–41.

[4] Blacking J. How musical is man? Seattle: University of Washington Press, 1973.

[5] Boltz MG. Tempo discrimination of musical patterns: effects due to pitch and rhythmic structure. Perception and Psychophysics 1998;60:1357–73.

[6] Boltz MG. The processing of temporal and nontemporal information in the remembering of event duration and musical structure. J Experimental Psychology: Human Perception and Performance 1998;24:1087–104.

[7] Boltz M, Jones MR. Does rule recursion make melodies easier to reproduce? If not, what does? Cognitive Psychology 1986;18:389–431.

[8] Brust JC. Music and language: musical alexia and agraphia. Brain 1980;103:367–92.

[9] Clarke EF. Rhythm and timing in music. In: Deutsch D, editor. The psychology of music. 2nd edition. San Diego, CA, USA: Academic Press, 1998.

[10] Deutsch D. The processing of structured and unstructured tonal sequences. Perception and Psychophysics 1980;28:381–9.

[11] Donchin E, Ritter W, McCallum WC. Cognitive psychophysiology: the endogenous components of the ERP. In: Callaway E, Tueting P, Koslow S, editors. Brain event-related potentials in man. New York, NY: Academic Press, 1978.

[12] Fasanaro AM, Spitaleri DL, Valiani R. Dissociation in musical reading: a musician affected by alexia without agraphia. Music Perception 1990;7:259–72.

[13] Fries W, Swihart A. Disturbance of rhythm sense following right hemisphere damage. Neuropsychologia 1990;28:1317–23.

[14] Griffiths TD, Johnsrude I, Dean JL, Green GR. A common neural substrate for the analysis of pitch and duration pattern in segmented sound? Neuroreport 1999;10:3825–30.

[15] Halpern AR. Organization in memory for familiar songs. Journal of Experimental Psychology: Learning, Memory, and Cognition 1984;10:496–512.

[16] Horikoshi T, Asari Y, Watanabe A, Nagaseki Y, Nukui H, Sasaki H, et al. Music alexia in a patient with mild pure alexia: disturbed visual perception of nonverbal meaningful figures. Cortex 1997;33:187–94.

[17] Kok A. Event related potentials (ERP) reflections of mental resources: a review and synthesis. Biological Psychology 1997;45:19–56.

[18] Krumhansl CL. Rhythm and pitch in music cognition. Psychological Bulletin 2000;126:159–79.

[19] Kutas M, Hillyard SA. Reading senseless sentences: brain potentials reflect semantic incongruity. Science 1980;207:203–5.

[20] Kutas M, Van Petten C, Besson M. Event-related potential asymmetries during the reading of sentences. Electroencephalogr and Clinical Neurophysiology 1988;69:218–33.

[21] Jones MR, Boltz M, Kidd G. Controlled attending as a function of melodic and temporal context. Perception and Psychophysics 1982;32:211–8.

[22] Judd T, Gardner H, Geschwind N. Alexia without agraphia in a composer. Brain 1983;106:435–57.

[23] MacLeod CM. Half a century of research on the Stroop effect: an integrative review. Psychological Bulletin 1991;109:163–203.

[24] Mavlov L. Amusia due to rhythm agnosia in a musician with LHD: a nonauditory supramodal defect. Cortex 1980;16:331–8.

[25] Miceli G, Capasso R. Calculation and number processing. In: Denes G, Pizzamiglio L, editors. Handbook of neuropsychology. Hove: Psychology Press, 1999.

[26] Monahan CB, Carterette EC. Pitch and duration as determinants of musical space. Music Perception 1985;3:1–32.

[27] Nittono H, Bito T, Hayashi M, Sakata S, Hori T. Event-related potential elicited by wrong terminal notes: effects of temporal disruption. Biological Psychology 2000;52:1–16.

[28] Palmer C, Kruhmhansl CL. Independent temporal and pitch structures in determination of musical phrases. Journal of Experimental Psychology: Human Perception and Performance 1987;13:116–26.

[29] Peretz I. Processing of local and global musical information by unilateral brain-damaged patients. Brain 1990;113:1185–205.

[30] Peretz I. Can we lose memory for music? A case of music agnosia in a nonmusician. Journal of Cognitive Neuroscience 1996;8:481–96.

[31] Peretz I, Kolinsky R. Boundaries of separability between melody and rhythm in music discrimination: a neuropsychological perspective. The Quarterly Journal of Experimental Psychology 1993;46A:301–25.

[32] Renault B, Ragot R, Lesevre N, Remond A. Onset and off set of brain events as indices of mental chronometry. Science 1982;215:1413–5.

[33] Ritter W, Simson R, Vaughan HG, Friedman D. Brain event related to the making a sensory discrimination. Science 1979;203:1358–61.

[34] Sadie S, editor. The new grove dictionary of music. New York: Macmillan, 1995.

[35] Schön D, Semenza C, Denes G. Oral reading of music. A selective deficit in one musical clef. Cortex 2001;37:407–21.

[36] Schönberg A. Harmonielehre. Wien: Universal Edition, 1922.

[37] Waters AJ, Underwood G, Findlay J. Studying expertise in music reading: use of a pattern matching paradigm. Perception and Psychophysics 1997;59:477–88.

[38] Waters AJ, Underwood G. Processing pitch and temporal structures in music reading: independent or interactive processing mechanisms? European Journal of Cognitive Psychology 1999;11:531–53.