Sight-reading expertise: Cross-modality integration investigated using eye tracking

Véronique Drai-Zerbib

Département de Psychologie, Université de Nice Sophia Antipolis, and LUTIN, Cité des sciences et de l'industrie de la Villette, France

Thierry Baccino

Université de Paris VIII, and LUTIN, Cité des sciences et de l'industrie de la Villette, France

Emmanuel Bigand

Université de Bourgogne, France

Abstract

It is often said that experienced musicians are capable of hearing what they read (and vice versa). This suggests that they are able to process and to integrate multimodal information. The present study investigates this issue with an eye-tracking technique. Two groups of musicians chosen on the basis of their level of expertise (experts, non-experts) had to read excerpts of poorly-known classical piano music and play them on a keyboard. The experiment was run in two consecutive phases during which each excerpt was (1) read without playing and (2) sight-read (read and played). In half the conditions, the participants heard the music before the reading phases. The excerpts contained suggested fingering of variable difficulty (difficult, easy, or no fingering). Analyses of first-pass fixation duration, second-pass fixation duration, probability of re-fixation, and playing mistakes validated the hypothesized modal independence of information among expert musicians as compared to non-experts. The results are discussed in terms of the processing cues and retrieval structures postulated by Ericsson and Kintsch (1995) in their model of expert memory.

Keywords

cross-modality, expertise, eye movements, psychology of music, sight-reading

Corresponding author:
Véronique Drai-Zerbib - Université de Nice Sophia Antipolis, Département de Psychologie, L.A.S.M.I.C., 24 Avenue des Diables Bleus, F-06357 Nice, France.
[email: Veronique.drai-zerbib@unice.fr]

Introduction

Music sight-reading

Music sight-reading consists of extracting visual information from a score to perform by simultaneous motor responses as playing or singing, laying on an auditory feedback. But it can also be done silently, without the mediation of instrument or voice. A common sense idea is that expert musicians can hear what they read from a score and visually represent the music that they are listening to. The composer Robert Schumann (Schumann & Schumann, 2009/1848) used it as a learning criterion. Skilled students were instructed to "hear music from the page" and create a mental representation of a piece after a single hearing, as if they had the music in front of them. Although this idea may seem obvious to most of us, and is generally thought to result from extensive playing of a musical instrument, there are still no satisfactory scientific explanations of this ability. How can one describe the transfer of information from one modality to the other by musicians of different levels of expertise, and how do they integrate musical information?

Cross-modality and expertise in music reading

The importance of inner hearing in musical sight-reading has been investigated using an interference paradigm, which showed that listening to distracting music affects inner hearing (Wöllner, Halfpenny, Ho, & Kurosawa, 2003). A number of studies have proposed explanations based on the notion of musical imagery[1] by analogy to the auditory imagery generated during silent reading of verbal material (inner speech). In particular, some authors postulate the existence of notational audiation, through which musicians access the musical representation of a piece of music by simply reading it (Gordon, 1993). According to Brodsky, Henik, Rubinstein, and Zorman (2003), the two major studies on auditory imagery (Reisberg, 1992) and musical imagery (Godoy & Jorgensen, 2001) do not provide any conclusive arguments on this topic. Using an embedded melody paradigm, Brodsky and colleagues (2003, 2008) showed that notational audiation relied heavily on kinaesthetic phonatory processes and pointed out "the profound reliance on phonatory and manual motor processing used during music reading". Furthermore, Fine, Berry, and Rosner (2006, p. 431) suggested an "increasing role for internal auditory representations with increasing expertise". However, they did not describe the exact nature of this audiation process, nor its underlying mechanisms. What about auditory, visual and motor imagery encoding? What about cross-modal integration? Is notational audiation activated by cues present in the musical score and, if so, how do the eye movements find out this information? Is it based on special memory capacities in individuals who possess this ability (i.e., experts)? Finally, notational audiation seems to be closely tied to singing, which involves a phonological production, but what about sight-reading on a keyboard where no verbal production is required?

Musical imagery and mental representation

Other empirical studies using a variety of methodologies (response times, event-related potentials [ERPs], etc.) have attempted to demonstrate cross-modality effects (not just musical), mostly by manipulating interference between different perception systems (olfaction, audition, vision, etc.; Holcomb & Anderson, 1993; Thompson & Paivio, 1994; Shimojo & Shams, 2001;

Luisa Dematté, Sanabria, & Spence, 2006; Robinson & Sloutsky, 2007). These studies can be classified into two categories, depending on their underlying hypothesis: (1) cross-modal conversion (recoding hypothesis), wherein information in one modality is converted into the other modality; and (2) cross-modal integration (amodal hypothesis), wherein information is not encoded in a modality-dependent way but is integrated at a higher level in an amodal representation. These two hypotheses are not mutually exclusive; they operate at different information-processing levels (perceptual for the former, conceptual for the latter) and depend on the individual's prior knowledge and skill level in the activity or task to be carried out. In our view, less expert musicians should use the recoding modality because they are more close to the written code (Drai-Zerbib & Baccino, 2005) while experts should be able to integrate the musical information into an amodal representation.

By bringing modal interference or ambiguous experimental situations into play, research based on the *recoding hypothesis* has provided evidence of a *conflict* between modalities (Shams, Kamitani, & Shimojo, 2000; Guttman, Gilroy, & Blake, 2005; Phillips-Silver & Trainor, 2007; Robinson & Sloutsky, 2007). For example, it is more difficult to recognize rhythmic sequences during encoding when incongruous auditory information is heard (Guttman et al., 2005). Similarly for motor signals, body movements during dancing can affect the auditory perception of rhythm (Philipps-Silver & Trainor, 2007). A simple auditory signal can also alter the perception of a visual stimulus and generate a visual illusion (Shams et al., 2000).

The *amodality hypothesis* is brought to bear in cases where a higher representation level must be accessed in order to integrate information from various sources. This integrating role is often assigned to memory representations and to the inference processes that build musical representations. Drai-Zerbib and Baccino (2005) investigated the role of expertise in musical sight-reading using an eye-tracking method. Expert and non-expert pianists first listened to a piano excerpt, read it, and then played it on a keyboard. Two versions of scores were used, with or without phrasing marks, either during the listening phase or the reading phase. They pointed out that the number and duration of fixations were significantly higher for non-experts compared to experts, particularly in no phrasing marks condition. Skilled musicians were found to have very low sensitivity to the written form of the score and seemed to reactivate a representation of the musical passage from the material listened to. In contrast, less skilled musicians were very dependent on the written code and on the input modality and must build a new representation based on visual cues. For language in particular, the amodality of semantic memory has been confirmed (Holcomb & Anderson, 1993; Rugg, Doyle, & Melan, 1993). Based on an ERP study using a semantic priming task, Holcomb and Anderson (1993) argued that amodal semantic representations are accessed by way of modality-specific encoding mechanisms. Further support for the amodality hypothesis has been obtained recently in brain imaging studies (functional Magnetic Resonance Imaging [fMRI], Magnetoencephalography [MEG]), where cerebral structures generally associated with different perceptual modalities were shown to be interconnected or overlap (Hasegawa et al., 2004; Just, Newman, Keller, McEleney, & Carpenter, 2004; Baumann et al., 2007).

The present study attempts to extend the idea of amodality, acknowledged for language and semantic memory, to the domain of music. Rather than speaking of auditory imagery – that seems inadequate in accounting for expert musicians' ability to hear what they read – this capacity will be seen as resulting from the *amodal* nature of the expert musician's memory. Memory is the basis not only for the mechanisms of visual encoding (Waters, Underwood, & Findlay, 1997; Waters, Townsend, & Underwood, 1998) and information retrieval (Gillman, Underwood, & Morehen, 2002), but also of inference-making processes (Lehmann & Ericsson, 1996). Saying that the expert musician's memory is amodal means that experts code musical

information independently of the input modality and can retrieve it regardless of how the information was perceived (visually or auditorily). It follows that perceptual cues might be less important for experts since they are capable of using their musical knowledge to compensate for missing or incorrect information. A more experienced performer may have both better representations of musical structure and better ability to apply these, or only one of these attributes. Conversely, less-expert musicians, who probably do not possess this ability can be assumed to go through a slower recoding phase. As a consequence we predict fewer fixations during reading for expert compared to non-expert. Furthermore, less-expert musicians can process even perceptual cues that might not be adapted for the execution since they seem very dependent of the written code (Drai-Zerbib & Baccino, 2005).

In this experiment, eye movements of expert and non-expert pianists were recorded as they sight-read piano excerpts, in order to (1) demonstrate the impact of a written code that supplies visuomotor cues (fingering);[2] and (2) validate the hypothesized modal independence of experts. Eye-tracking measurements have been largely used in psycholinguistic studies to investigate the cognitive processes that intervene during the reading of a text (Tinker, 1946; Kennedy, Murray, O'Regan, & Levy-Schoen, 1987; Morris, Rayner, & Pollatsek, 1990; Rayner, 1993, 1998; Reichle, Pollatsek, Fisher, & Rayner, 1998; Chace, Rayner, & Well, 2005). The method allows distinguishing early perceptual processes involved during encoding from late cognitive processes entailed by integrative operations. The duration of eye fixations and speed of eye saccades highlight the processes that govern reading and integration (Baccino, 2002). Some studies have compared eye fixations in reading language and music (Weaver, 1943; Goolsby, 1994; Kinsler & Carpenter, 1995; Rayner & Pollatsek, 1997). As a main outcome, reading experts in music were shown to differ from novice readers by their number, the place and the duration of their eyes fixations on the score (Jacobsen, 1942; Weaver, 1943; Truitt, Clifton, Pollatsek, & Rayner, 1997; Waters et al., 1997). The pattern of eye movements "can be a sensitive indicator of cognitive operations while reading music" (Rayner & Pollatsek, 1997, p. 50). So this method furthers the understanding of multimodal information processes and crossing modal integration. In the present study, an auditory representation was activated before the music was read in half of the experimental conditions. Fingering was manipulated in terms of playing difficulty: easy, difficult, or no fingering, and constituted a visuomotor cue. This study should allow us to distinguish perceptive, mnemonic and motor involvements in musical reading according to the level of experts. According to the amodal hypothesis, experts should exhibit a stable pattern of eye fixations during the reading and the sight-reading task. This stability may be shown by the same number and duration of fixations, no matter what type of information was provided (auditory, visual, or visuo-motor). Non-experts, however, should be affected by the different modality of the musical information perceived, as suggested by the recoding hypothesis.

Method

Participants

The 25 participants were piano students or teachers at the National Conservatory of Music in Nice, France. They were divided into two expertise groups on the basis of their piano-playing skill. It should be pointed out here that musical knowledge expertise was not separated from the playing skill. There were 15 experts (piano teachers, accompanists, or students who had already obtained their degree) with more than 12 years of practice, and 10 non-experts (piano students who had been studying at the conservatory for six to eight years).

Musical material

The musical material consisted of 36 piano excerpts, each four measures long. The excerpts were taken from the classical tonal repertoire. A listing of musical excerpts and their composer can be found in Appendix 1.

Three versions of each excerpt (12 per fingering condition) were generated according to whether the fingering was given or not (control condition). On the fingered excerpts, fingering was noted on measures two and three by a sight-reading teacher from the conservatory. Two types of fingering were defined, one difficult and one easy. Difficulty was rated by three separate experts on a Likert scale ranging from 1 (very easy) to 5 (very difficult). The difference of means between the two fingering difficulty levels was significant (difficult fingering = 3.78, easy fingering = 1.99, $t(70) = 13.36$, $p < .001$). Fingering was considered easy when it corresponded to that used logically to play the piece, and difficult (but not incorrect) when it involved crossing over or finger positions that were hard to achieve and not suited to the piece. In piano playing, the fingers are numbered from 1 to 5 (from the thumb to the fifth finger). The presupposed optimal fingering always involves a trade-off between physiological, anatomical, cognitive, and interpretational constraints, but there is no absolute optimal fingering (Parncutt, Sloboda, Clarke, Raekallio, & Desain, 1997). Good fingering should allow the musician to play a passage as comfortably as possible. Figure 1 gives an example of fingering for an excerpt.

The musical excerpts were input using Final™ software and saved in .bmp format. For the auditory phase, the excerpts were played on a Steinway™ grand piano by a piano teacher from the conservatory and recorded using a microphone and a Sony™ minidisk recorder. The files were processed using Sound Forge™ software and then saved in .wav format.

Apparatus

The ocular data was sampled at a frequency of 50 Hz (every 20 ms) using the TOBII Technology 1750™ eye-tracking system connected to two portable PC computers (Dell Latitude D505™ and Dell Latitude D800™). The music staves were displayed on a 17″ screen at an image resolution of 1024 × 768 pixels. The excerpts were heard through Plantronics™ headphones. The participants sight-read the excerpts on a Yamaha P120™ 88-key electronic piano and their
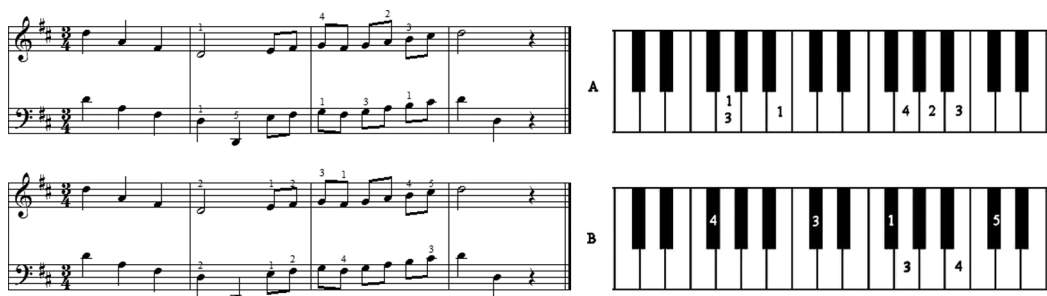

Figure 1. Diagram of fingering (measure 3 only) on an excerpt from Haydn's Menuetto and Aria. (A) Easy fingering; (B) difficult fingering. The numbers on the keyboard refer respectively to the following fingers 1 = thumb ; 2 = index; 3 = middle; 4 = ring finger; 5 = little finger.

playing was recorded on a portable Maxdata™ PC computer via the MIDI Evolution EV10™ interface and Sound Studio II™ software. The experiment was programmed in E-prime™ software.

Experimental procedure

After the eye tracker was calibrated (nine-point procedure), a practice phase on three excerpts was run and then followed by the experiment proper. The experimental phase included the following steps (see Figure 2): (1) a fixation cross was displayed on the left side of the screen in front of where the excerpt would appear and the participant either heard or did not hear a musical excerpt being played on the piano; (2) the participant read the corresponding musical score for the first time (discovery) and informed the experimenter when the reading was completed (this was the signal for the experimenter to press the space bar to go on to the next step); and (3) the participant reread the same excerpt while playing it on a MIDI keyboard (sight-reading). After the music was played, the experimenter pressed the space bar again to display the next excerpt. This procedure was repeated for the 36 excerpts that had been randomized across participants. The participant's eye movements and musical production were recorded each time.

Experimental design

The experimental design was a quasi-complete repeated measures design with one between-subject factor, expertise (15 expert pianists, 10 non-expert pianists), and three within-subject
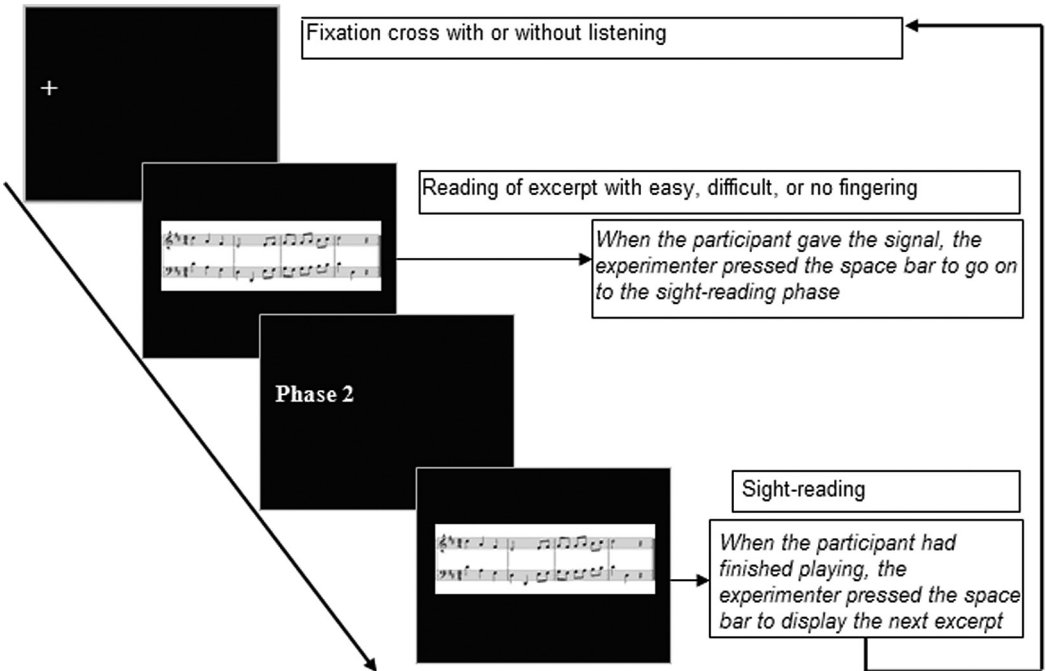

Figure 2. Diagram of the experimental procedure.

Figure 3. Breakdown of the excerpt into nine areas of interest: clef area, upper staff (right hand measures r1 to r4), lower staff (left hand measures l1 to l4).

factors, reading phase (reading alone, sight-reading), preliminary listening (with/without listening), and fingering (difficult, easy, no fingering). This made for a total of 12 experimental conditions counterbalanced across participants.

Each participant saw all 36 excerpts, three per experimental condition. The eye-tracking data was based on nine areas of interest (AOIs) in the musical staff: the clef area (containing the treble and bass clefs, the key signature, and the time signature), the four treble clef measures played with the right hand (r1 to r4), and the four bass clef measures played with the left hand (l1 to l4). This added another within-subject factor to the design: the area of interest (AoI). Figure 3 shows the AoIs outlined on a sample staff.

Four dependent variables were analyzed for each AoI. First-pass fixation duration (F1D) was the mean duration of initial fixation entering into an AoI (the mean of all fixations between the first entry into the AoI and the first exit from the AoI). Second-pass fixation was the mean duration to look back an AoI similar to rereading (the mean of all fixations occurring after first pass fixation)(Rayner, 1998). The probability of re-fixation (PR) was the probability to look back at an AoI. Playing errors were incorrect or missing notes played during the participants' execution on the keyboard; errors were counted in the recorded MIDI files.

Results

Analyses of variance were conducted on the four dependent variables, and we used an alpha < .05 as a significance level. All results of the analysis of variance were Greenhouse and Geisser corrected. Post-hoc Scheffé tests were used to evaluate the significance. Table 1 summarizes the mean results obtained for the whole score.

Mean first-pass fixation duration (F1D)

For all participants pooled, F1D varied significantly across the AoIs of the excerpt, $F(8, 184) = 22.38$, $\eta^2 = .49$, $p < .001$. The clef area was fixated less than the other areas, $F(1, 23) = 86.69$, $p < .001$, and the measures on the treble clef or upper staff (played with the right hand) were fixated more than the bass clef or lower staff measures (played with the left hand), $F(1, 23) = 34.67$, $p < .001$. The interaction between AoI and excerpt listening is particularly interesting, $F(8, 184) = 39.96$, $\eta^2 = .63$, $p < .001$. F1D was longer on the upper staff only when the musicians had not heard the music first, $F(1, 23) = 110.80$, $p < .001$. This suggests

Table 1. Mean first-pass fixation duration (F1D), mean second-pass fixation time (F2D), probability of re-fixation (PR), and number of mistakes, by expertise level (expert, non-expert), reading phase (except for mistakes) (L1: reading alone, L2: sight-reading), fingering (difficult, easy, no fingering), and listening (with, without) for the whole score and all subjects.

		Difficult fingering		Easy fingering		No fingering	
		Listening	No listening	Listening	No listening	Listening	No listening
F1D							
Experts	L1	306 (22.6)	324 (23.6)	292 (22.5)	332 (19)	334 (26.2)	286 (23.6)
	L2	318 (18)	333 (20.5)	310 (19.4)	322 (22.6)	336 (24.8)	319 (19.8)
Non-experts	L1	382 (27.8)	377 (28.9)	337 (27.5)	380 (23.7)	373 (32.1)	345 (28.9)
	L2	345 (22.1)	370 (25)	397 (23.8)	354 (27.6)	382 (30.4)	355 (24.3)
F2D							
Experts	L1	568 (93.6)	609 (91.4)	506 (95.7)	577 (83.1)	571 (87.3)	522 (13.9)
	L2	482 (56.3)	501 (52.5)	450 (67)	501 (52.5)	555 (69.6)	520 (64.8)
Non-experts	L1	866 (114.6)	945 (111.9)	912 (117.2)	952 (101.8)	997 (106.9)	998 (140)
	L2	699 (68.9)	658 (76.1)	683 (82)	631 (64.3)	701 (85.2)	681 (79.3)
PR							
Experts	L1	0.46 (0.04)	0.52 (0.04)	0.52 (0.05)	0.56 (0.04)	0.59 (0.05)	0.55 (0.05)
	L2	0.54 (0.04)	0.54 (0.05)	0.59 (0.05)	0.57 (0.05)	0.64 (0.04)	0.61 (0.05)
Non-experts	L1	0.69 (0.05)	0.77 (0.05)	0.77 (0.06)	0.73 (0.05)	0.76 (0.06)	0.77 (0.05)
	L2	0.60 (0.05)	0.69 (0.06)	0.70 (0.06)	0.67 (0.06)	0.73 (0.05)	0.71 (0.06)
Mistakes							
Experts		2.58 (1.56)	1.63 (1.27)	2.24 (1.33)	3.51 (2.13)	3.62 (1.36)	1.49 (1.16)
Non-experts		7.23 (6.40)	5.79 (4.67)	5.77 (3.38)	7.77 (5.19)	8.4 (6.66)	4.73 (4.51)

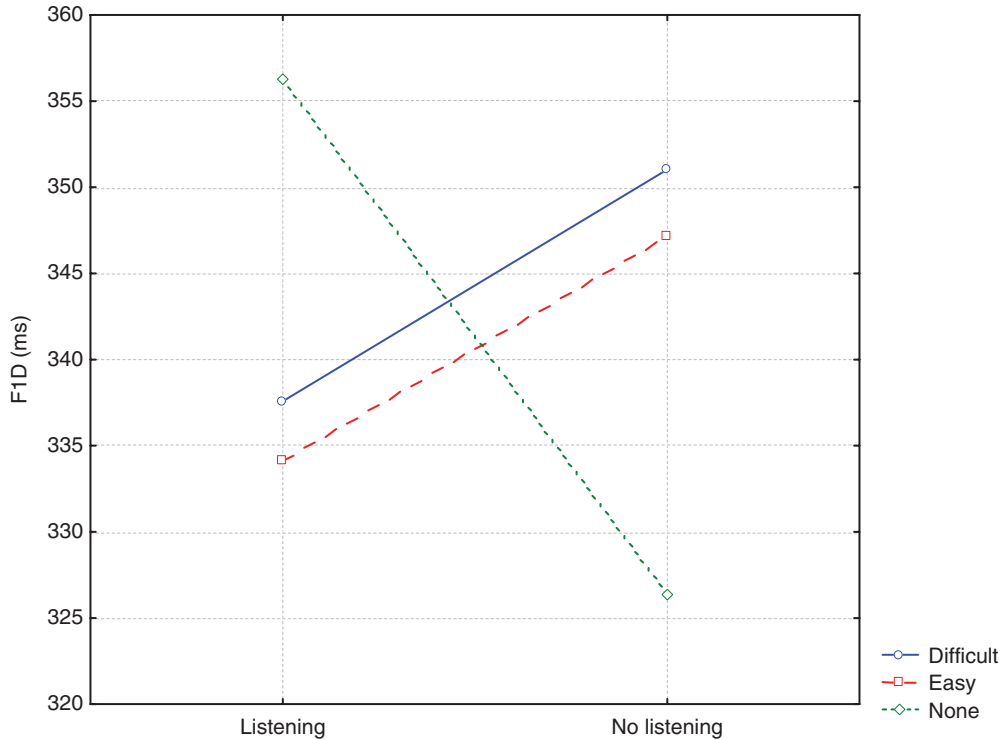*Note.* Standard deviations are shown in parentheses.

Figure 4. Mean F1D as a function of listening condition (with/without listening) and fingering (difficult, easy, no fingering). Y scale has been dimensioned as function of the observed data.

that hearing the music allowed the pianists to memorize the treble clef notes and this facilitated later reading. The interaction obtained between listening and fingering (Figure 4) argues in favor of cross-modal integration, $F(2, 46) = 11.31$, $\eta^2 = .33$, $p < .001$. When easy or difficult fingering was noted on the staff, preliminary listening lowered F1D compared to the no-listening condition, $F(1, 23) = 5.56$, $p < .05$. The opposite occurred when there was no fingering, in which case initial listening caused a higher F1D than in the no-listening condition, $F(1, 23) = 13.83$, $p = .001$. This suggests that reading the fingered music had a greater facilitating effect for the pianists when they had an auditory reference.

The significant three-way interaction between reading, listening, and fingering (Figure 5) showed that fingering had an impact when the music was heard first (the difference between the two fingering conditions was non-significant when there was no listening), $F(2, 46) = 6.07$, $\eta^2 = .21$, $p < .01$. On the reading alone phase, F1D was longer for difficult fingering than for easy fingering, $F(1, 23) = 7.53$, $p < .025$. The opposite was found for the sight-reading phase, where F1D was shorter for difficult fingering than for easy fingering, $F(1, 23) = 6.88$, $p < .025$. This can be explained by considering that hearing the music in advance pre-activated or pre-selected the necessary fingering for playing the excerpt. On the first reading, the difficult fingering was judged unsuitable so the musicians looked longer at the music in an attempt to solve this problem. Once the difficult fingering problem was solved, the second reading was faster. The greater F1Ds on the first reading than on the second argues in favor of a *trade-off* strategy consisting of processing difficulties on the first reading in order to facilitate later execution.
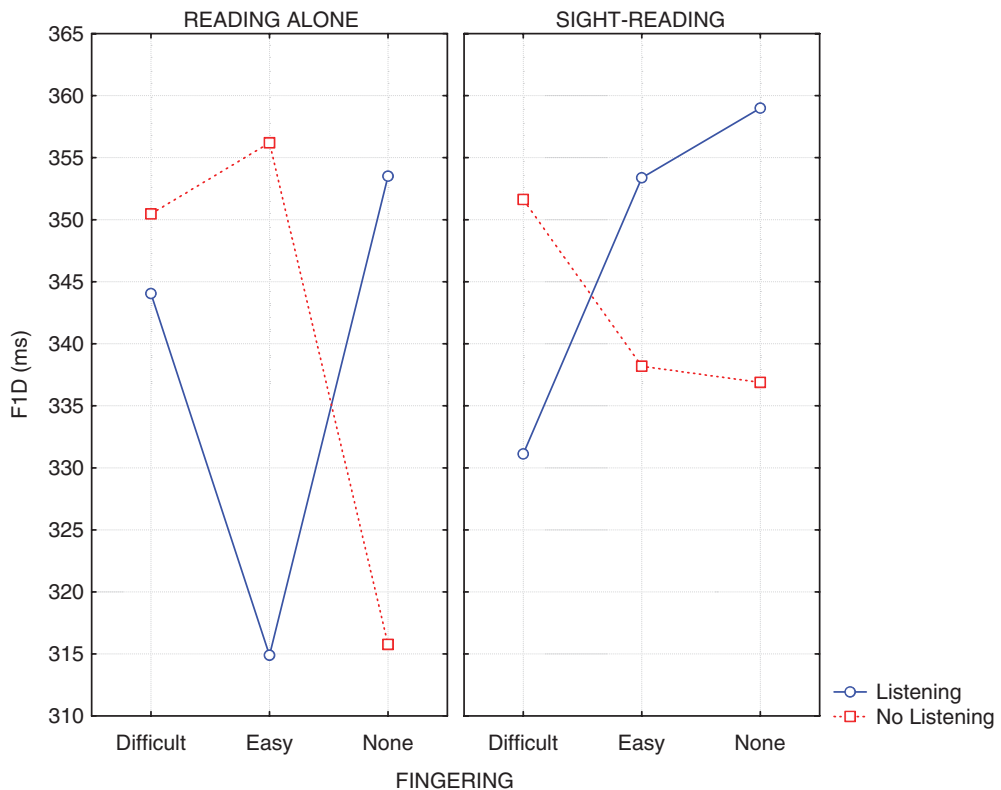
Figure 5. Mean F1D on the first and second readings (reading alone, sight-reading), as a function of listening condition (with/without listening) and fingering (difficult, easy, no fingering). Y scale has been dimensioned as function of the observed data.

Although there was no main expertise effect, the impact of expertise showed up in the four-way interaction between expertise, listening, fingering, and AoI, $F(16, 368) = 2.54$, $p < .001$, $\eta^2 = .10$. A processing difference across piano-playing expertise levels was observed in the fingered-measure comparisons (measures r2, r3, l2, and l3). When the pianists were experts and the fingering was difficult, F1D was shorter with listening than without, $F(1, 23) = 10.15$, $p < .01$; yet when the fingering was easy, F1D did not vary significantly; and when there was no fingering, F1D was longer with listening than without, $F(1, 23) = 6.17$, $p < .05$. For non-expert pianists, none of these comparisons were significant. It thus seems that only the experts did not process the difficult fingering when they had heard the music first.

Mean second-pass fixation time (F2D)

While the experts had shorter F2Ds than the non-experts, $F(1, 23) = 8.91$, $p < .01$, the behavior of the two groups was similar, no matter what AoI was considered, as shown by the significant interaction between expertise and AoI, $F(8, 184) = 3.80$, $\eta^2 = .14$, $p < .001$. Except for different second-pass fixation times on the first measure, where only non-experts took more

time than on the other measures, $F(1, 23) = 20.99$, $p < .001$, F2D was higher on the upper staff measures than on the lower staff ones for all musicians, $F(1, 23) = 47.43$, $p < .001$. Moreover, as found for F1D, F2D on the clef area was significantly shorter than on the other AoIs, $F(1, 23) = 13.62$, $p < .01$. A significant interaction between AoI and listening, $F(8, 184) = 20.93$, $\eta^2 = .48$, $p < .001$, indicated that musicians looked back at the clef area more often in the initial-listening condition, $F(1, 23) = 33.50$, $p < .001$. This listening effect on clef-area fixations also varied with fingering difficulty: when no fingering was present on the staff, second-pass fixation time was greater than when difficult fingering was shown, $F(1, 23) = 5.11$, $p < .05$, and it was even greater for easy fingering, $F(1, 23) = 11.64$, $p < .01$. Only when the music had not been heard in advance did the musicians look back more frequently at difficult fingering measures than at easy fingering ones, $F(1, 23) = 20.02$, $p < .01$.

Fingering had a main effect on F2D, $F(2, 46) = 3.83$, $\eta^2 = .14$, $p < .05$. Contrary to our predictions, however, the presence of fingering of either difficulty level triggered shorter F2Ds than the absence of fingering, $F(1, 23) = 7.02$, $p = .01$. There was a significant interaction between fingering and AoI, ($F(16, 368) = 4.67$, $\eta^2 = .17$, $p < .001$), and between fingering, AoI, and expertise, $F(16, 368) = 2.28$, $\eta^2 = .09$, $p < .01$. As stated earlier, second-pass fixations were longer without fingering than in all other cases, but this effect was greater among expert pianists than among non-experts, $F(1, 23) = 6.32$, $p < .025$.

F2D was significantly higher during reading alone than during sight-reading, $F(1, 23) = 4.64$, $\eta^2 = .17$, $p < .05$. Logically, this means that sight-reading was facilitated by the initial reading. But it also suggests that on the first reading, the musicians engaged in a motor-planning process, which they could not do on the second reading because they were forced to play. This explanation would be coherent with findings from Goebl and Palmer (2006) that showed anticipatory movements of pianists' fingers. This anticipation would take place during this first reading. The interaction between reading and listening, $F(1, 23) = 5.27$, $\eta^2 = .19$, $p < .05$, is of particular interest to the issue of cross-modality insofar as it shows that initial listening had an impact on the first reading only, where it triggered shorter fixations, $F(1, 23) = 4.50$, $p < .05$, whereas no difference was found on the second reading, $F = 1.17$, ns.

The three-way interaction between fingering, reading, and expertise (Figure 6), $F(2, 46) = 4.08$, $\eta^2 = .15$, $p < .025$, showed for all fingering conditions that the experts' second-pass fixations did not differ significantly across readings, $F < 1$, whereas for the non-experts, they were higher on the first reading, $F(1, 23) = 5.26$, $p < .05$. In fact, non-experts obtained higher F2Ds than experts solely for reading alone (386 ms), $F(1, 23) = 7.39$, $p < .025$. For sight-reading, the difference between the two levels of expertise (174 ms) was non-significant, $F < 1$.

The three-way interaction between reading, AoI, and expertise, $F(8, 184) = 2.99$, $\eta^2 = .21$, $p < .01$, confirmed that experts differed from non-experts only on the first reading, $F(1, 23) = 7.39$, $p < .025$. On the sight-reading phase, there was no difference between the two groups of pianists. Apparently, the non-experts took longer second-pass looks at the excerpts during the first reading in order to prepare for a more efficient motor execution (Goebl & Palmer, 2006).

Probability of re-fixation (PR)

The expert pianists looked back at the staff significantly less often than the non-experts did, $F(1, 23) = 8.48$, $\eta^2 = .27$, $p < .01$. Probability of re-fixation (PR) differed significantly across fingerings too (difficult < easy < none), $F(2, 46) = 27.86$, $\eta^2 = .55$, $p < .001$. Moreover, it varied
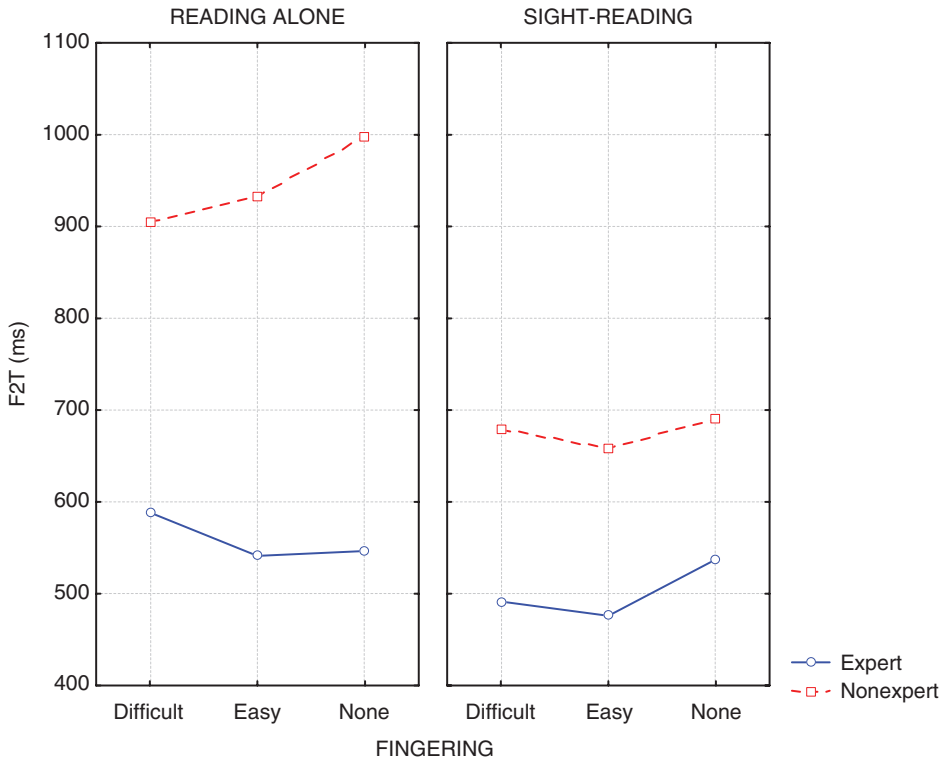
Figure 6. Mean F2D on the first and second readings (reading alone, sight-reading), as a function of fingering (difficult, easy, no fingering) and expertise (expert, nonexpert). Y scale has been dimensioned as function of the observed data.

as a function of AoI in exactly the same way as F1D and F2D did; that is, there was far less probability of re-fixation on the clef area than on the other AoIs, especially during reading alone, $F(1, 23) = 59.56$, $p < .001$, and there was less probability of re-fixation on the lower staff than on the upper staff, $F(8, 184) = 39.24$, $p < .001$.

An interaction between listening and fingering was also obtained, $F(2, 46) = 10.67$, $\eta^2 = .32$, $p < .001$. When the fingering was difficult, the pianists looked back at the staff less often when they had heard the music first, $F(1, 23) = 22.68$, $p < .001$, while for the other fingerings, the listening effect was non-significant, $F < 1$. The impact of listening in the presence of difficult fingering was especially great for the non-expert pianists, $F(1, 23) = 21.03$, $p < .001$. This suggests that, for non-experts, listening appears very helpful for solving the difficult fingerings. Conversely, for the experts, the probability of re-fixation was higher with listening than without when there was no fingering on the staff, $F(1, 23) = 4.32$, $p < .05$.

Correlations between fixation time and number of errors

The MIDI files were used to count the number of errors made by the musicians as they sight-read on the keyboard (see Table 1). The analysis of variance on the mistakes yielded the expected effect of expertise, $F(1, 19) = 8.65$, $\eta^2 = .31$, $p < .01$; that is, experts made significantly fewer

Table 2. Correlations (Bravais Pearson's r) between total fixation duration and number of playing mistakes, by reading phase (L1: reading alone, L2: sight-reading) and expertise (experts, non-experts)

		Experts (*N* = 13)	Non-experts (*N* = 8)
	L1	**0.57 ***	0.33
	L2	0.38	**0.86 ****
L1	Listening	0.52	0.33
	No listening	**0.59 ***	0.33
	Difficult fingering	**0.57 ***	0.30
	Easy fingering	0.52	0.49
	No fingering	**0.60 ***	0.20
L2	Listening	0.37	**0.85 ****
	No listening	0.33	**0.87 ****
	Difficult fingering	0.35	**0.82 ***
	Easy fingering	0.48	**0.86 ****
	No fingering	0.22	**0.88 ****

Note. Significant correlations are shown in bold. *p < .05; **p < .01.

mistakes than non-experts did. To assess the impact of the ocular intake of information during sight-reading, correlations were computed between the total fixation time (F1D + F2D) and the number of mistakes made during playing (incorrect notes played; Table 2).

Total fixation time was positively correlated with the number of playing mistakes, but only for reading-alone fixation time for experts ($r(13) = 0.57$, $p < .05$), and only for sight-reading fixation time for non-experts ($r(13) = 0.86$, $p < .01$). These correlations mean that the longer the fixation time, the greater the number of playing mistakes, but the interesting point is that this did not happen at the same time for expert and non-expert pianists. The experts made a greater number of mistakes when they had spent more time looking at the staff during the first reading, particularly when the fingering was absent ($r(13) = 0.60$, $p < .05$) or difficult ($r(13) = 0.57$, $p < .05$) and when there was no listening phase ($r(13) = 0.59$, $p < .05$). During reading alone (first reading), then, experts seem to have anticipated the motor difficulties (difficult fingering or no fingering) they would encounter during sight-reading (second reading). This points to an interesting cross-modality effect: the mere viewing of the music facilitated preparation for motor execution. The fact that experts solved the motor-planning problems on the first reading could account for why there was no longer a significant correlation on the second reading for these pianists. Conversely, the non-experts seem to have been unable to perform this motor anticipation during the first reading, so they discovered the motor difficulties (fingering, etc.) while they were sight-reading.

Discussion

First, this experiment pointed out a number of behavioral constants in the initial processing of musical scores by pianists (first-pass fixations), irrespective of their level of expertise. The upper staff (right hand) was processed more than the lower staff (left hand) except when the music was heard before being read. In the latter case, the lack of a difference between the right and left hands suggests that the musicians processed the upper staff during prior listening and this facilitated later processing during reading. The clef area, a crucial area that determines the key

of the piece, was read faster than the other areas, but paradoxically, the pianists looked back at this zone more when they had heard the music first. This can be explained in terms of the idea that the musicians built a model of the music's key while listening, and then validated that model during the first few fixations on the clef area. Weaver (1943) showed that the structure of a piece of music partially determines how a musician will read the notes on the page. As a general rule, one can assume that an initial hearing facilitates recognition of the music, but our results also suggest that preliminary listening also helps pianists plan their motor execution (finger positions) for later playing.

Second, in our study, the experts did not process the difficult fingerings on the first pass when they had heard the piece in advance, whereas on the second pass they looked more at the difficult fingerings than at the easy ones when they hadn't heard the piece. Thus their reading of the fingered staves was more efficient on the first reading when preliminary listening was possible. Recall that the difficult fingerings noted on the staves were not "incorrect" but difficult to execute. It seems logical that such fingerings would be ignored during actual playing, especially by musicians who were experts and who had listened to the music in advance. This is consistent with the results of Clarke, Parncutt, Raekallio, and Sloboda's (1997) study on nine professional pianists. These authors showed that, on tasks requiring allocation of attentional resources such as sight-reading or performing in public, expert musicians tend to adopt the standard fingering they were taught when learning to play the instrument (while we think that there are no general best or "good fingerings" across different pianists). Parncutt and colleagues (1997) also stressed that expert pianists are inclined to use the same fingerings for recurring passages in a given piece, and also to transfer to new pieces familiar fingering they have already figured out on similar pieces. An overall explanation may stem from the fact that becoming an instrumental musician involves starting early and practicing regularly, for several hours a day (Ericsson, Krampe, & Tesch-Römer, 1993; Ericsson & Lehmann, 1996). Music performance is a sensorimotor activity that requires precise timing of several hierarchically-organized actions implemented by various effectors, in accordance with the particular instrument (Zatorre, Chen, & Penhune, 2007). Several studies using cerebral imagery showed that the same cortical areas were activated whether piano music was being read or played (Meister et al., 2004), which is consistent with the idea that music reading involves a sensorimotor transcription of the music's spatial code (Stewart et al., 2003). One of the sensorimotor activities taking place during music reading on the piano consists of anticipating the positions of the fingers. The presence of fingering on the music aids in this anticipation process by providing visuomotor cues, and helps the pianist to find the fingering combinations prescribed for virtuoso playing. So the fingering is an important visuo-motor cue that allows the musician to anticipate the position of his fingers on the instrument (Parncutt et al., 1997).

Another interesting finding for the experts was their tendency to make use of the preliminary listening phase to determine the suitability of the proposed fingering. These pianists looked back at the staves without fingering more when they had heard the music first, whereas the non-experts focused on the excerpts containing difficult fingering. The experts probably rapidly assessed unsuitable fingering and chose to ignore it during sight-reading, whereas the non-experts (less inclined to question the proposed musical notation) attempted to resolve these difficult fingering problems. Clearly the ability to rapidly access proper fingering is acquired through extensive practice, during which the musician learns to quickly decide what fingering combinations should be used to play a given piece. It has been shown, for example, that pianists avoid using the fourth or fifth finger on a black key (Sloboda, Clarke, Parncutt, & Raekallio, 1998). In our experiment, initial listening seems to have been critical in helping the experts quickly select suitable fingering. They modulated their visual information intake by rapidly

scanning the staves showing inappropriate (difficult) fingering and returning more to the ones with no fingering noted. Musical expertise seems to be characterized by considerable eye-fixation flexibility and adaptation to the difficulty of the material being read. This behavior is similar to that observed for text reading (Rayner, 1998).

The lack of an initial processing difference between the two groups of musicians can be accounted for by the skill level of our non-expert pianists: they averaged seven years of musical instruction, so they were not beginners. However, the effect of musical expertise showed up quite clearly on the re-processing of the musical scores. Whether on second-pass fixation time or probability, the experts looked back significantly less at the music than the non-experts did. Expertise-dependent oculometric patterns have already been found for text reading (Rounds, Manley, & Norris, 1991), but the present study is the first to observe this for music reading. We found in particular that non-experts took longer second looks at the music during the first reading, which suggests that they had a greater need for motor planning than experts did. This longer reading time turned out to be beneficial, to the extent that non-experts read in a similar way to experts on the second reading.

Finally, the correlation observed here between visual information intake and the number of mistakes made as the pieces were being played is clearly indicative of the experts' ability to take advantage of the first reading in order to anticipate potential motor problems (difficult fingering) they might have during sight-reading (second reading). This finding supplies an important argument in support of the hypothesized cross-modal capacities of expert memory. For experts, the mere viewing of a musical score may facilitate planning and preparation for motor execution, whereas non-experts do not appear to have this cross-modal integration ability. Our non-experts seem to have discovered the difficulties as they were playing, so the preliminary reading had no impact on execution.

As a whole, these findings argue in favor of efficient cross-modal integration among expert musicians, and they extend clearly our past results (Drai-Zerbib & Baccino, 2005). Musical expertise can be defined as the ability to break away from the written code and to free oneself not only of the particular information intake modality, here visual or auditory (Brodsky et al., 2003; Drai-Zerbib & Baccino, 2005; Yumoto et al., 2005), but also of visuomotor information (fingering), as our study showed. Through lengthy training in sight-reading and years of playing, the expert is able to retrieve from memory the necessary information for efficiently processing the musical material (Ericsson et al., 1993). Intensive practice enables the musician not only to coordinate visual information intake (eye movements) and motor behavior (playing), as many studies on eye-hand span have shown (Sloboda, 1974; Rayner & Pollatsek, 1997; Truitt et al., 1997), but also to use auditory feedback to control and anticipate motor execution in accordance with the desired musical style. Thus, expertise in sight-reading, as in all complex activities, relies heavily on memory structures. Sloboda (1984) stressed that expert sight-reading is largely the result of a good visual memory for musical notation and our findings showed that expert musicians extract the relevant information more efficiently before translating it into motor commands.

A model of expert memory that can explain our results is found in Ericsson and Kintsch's (1995) theory of long-term working memory. This model suggests that experts organize their knowledge into so-called "retrieval structures" in long-term memory (LTM) in order to develop efficient retrieval strategies that surpass short-term memory capacities (STM) (Chase & Simon, 1973; Ericsson & Kintsch, 1995; Gobet, 1998). Based on this model, one can assume that musical knowledge structures are directly activated by visual, auditory, and motor retrieval cues (Williamon & Egner, 2004). Several studies have suggested that experts use hierarchical retrieval systems to recall encoded information (Halpern & Bower, 1982; Chaffin & Imreh, 1997; Aiello, 2001; Williamon & Valentine, 2002; Drai-Zerbib & Baccino, 2005; Chaffin, 2006). Two empirical

studies with an experienced pianist (Chaffin & Imreh, 1997; Chaffin, 2006) showed that this musician used the structure of the music to organize execution and memorize new pieces with only a few hours of practice. She relied on cues to make LTM memory retrieval more efficient, organized her performance of a piece around its formal structure, and stopped at the boundary of musical phrases rather than in the middle of a section. In the same vein, Aiello (2001) showed that classical concert pianists memorize a work they have to play by analyzing the musical score in detail and taking more notes about the musical elements of the piece than non-experts, who simply learn the piece by heart without analyzing it. Accordingly, expert musicians rapidly index and categorize musical information to form meaningful units (Halpern & Bower, 1982) which they use later when practicing or during a performance (Williamon & Valentine, 2002). Thanks to the use of organized knowledge structures stored in memory, expert pianists can reconstruct information missing from the musical score. Drai-Zerbib and Baccino (2005) found that, compared to non-experts, expert musicians were less tied to musical notation since they were capable of reinserting phrasing into music from which it had been removed, and that the experts' relative written-code independence increased when they were provided with an auditory rendition of the piece first. It would seem that expert musicians are able to construct an amodal representation of musical phrases; that is, they have the ability to represent a work regardless of whether the input modality is visual or auditory. This hypothesis is corroborated by other studies showing that, when a musical excerpt is presented for reading, the auditory imagery of the future sound is activated (Yumoto et al., 2005). Further research should address the question of how this amodal kind of memory representation develops in experts, for this capacity is the keystone of musical expertise.

Conclusion

In brief, this paper attempts to demonstrate that more experienced performers are better able to transfer learning from one modality to another, which can be in support of theoretical work by Ericsson and Kintsch (1995): more experienced performers better integrate knowledge across modalities. This view relies on the general flexibility shown in the experts' behavior. They had flexibility in the choice of fingerings: experts ignored unsuitable fingerings during sight-reading because they had processed them in an earlier reading and judged their inconsistency. Additionally, they had flexibility in the visual information intake – experts optimally used the first reading (more efficient fixations) to anticipate potential playing problems that could occur later. This flexibility is based on their own musical knowledge that compensates for the inconsistency of some situations and allows in performing better by anticipating difficulties. Flexibility might be an indicator of fluent sight-reading and it would be interesting in a further practical application to elaborate reading tests (based on the type of factors used here: difficult fingerings, listening before reading) that might assess the level of expertise.

Notes

1. Musical imagery is the mental capacity for imagining music inside the head, in the absence of a directly audible sound source (recall, or invent musical sound through "inner ear").
2. Fingering is represented with numbers written above each note that tell the player which finger to use.

References

Aiello, R. (2001). Playing the piano by heart, from behavior to cognition. *Annals of the New York Academy of Science*, *930*, 389–393.

Baccino, T. (2002). Oculométrie cognitive [cognitive eye tracking]. In G. Tiberghien (Ed.), *Dictionnaire des sciences cognitives* [Dictionary of cognitive science] (pp. 100–101). Paris: Armand Colin.

Baumann, S., Koeneke, S., Schmidt, C. F., Meyer, M., Lutz, K., & Jancke, L. (2007). A network for audio-motor coordination in skilled pianists and non-musicians. *Brain Research*, *1161*, 65–78.

Brodsky, W., Henik, A., Rubinstein, B. -S., & Zorman, M. (2003). Auditory imagery from musical notation in expert musicians. *Perception & Psychophysics*, *65*(4), 602–612.

Brodsky, W., Kessler, Y., Rubinstein, B., Ginsborg, J., & Henik, A. (2008). The mental representation of music notation: Notational audiation. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(2), 427–445.

Chace, K. H., Rayner, K., & Well, A. D. (2005). Eye movements and phonological parafoveal preview: Effects of reading skill. *Canadian Journal of Experimental Psychology*, *59*(3), 209–217.

Chaffin, R. (2006). Learning Clair de Lune: Retrieval practice and expert memorization. *Music Perception*, *42*(4), 377–393.

Chaffin, R., & Imreh, G. (1997). "Pulling teeth and torture": Musical memory and problem solving. *Thinking & Reasoning*, *3*(4), 315–336.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, *4*(1), 55–81.

Clarke, E. F., Parncutt, R., Raekallio, M., & Sloboda, J. A. (1997). Talking fingers: An interview study of pianists' views on fingering. *Musicae Scientiae: TheJjournal of the European Society for the Cognitive Sciences of Music*, *1*(spring), 87–107.

Drai-Zerbib, V., & Baccino, T. (2005). L'expertise dans la lecture musicale: Intégration intermodale [Expertise in reading music: Intermodal integration]. *L'Année Psychologique*, *105*, 387–422.

Ericsson, K. A, & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, *102*, 211–245.

Ericsson, K. A., Krampe, R., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, *100*(3), 363–406.

Ericsson, K. A., & Lehmann, A. C. (1996). Expert and exceptional performance: Evidence of maximal adaptation to task constraints. *Annual Reviews of Psychology*, *47*, 273–305.

Fine, P., Berry, A., & Rosner, B. (2006). The effect of pattern recognition and tonal predictability on sight-singing ability. *Psychology of Music*, *34*(4), 431–447.

Gillman, E., Underwood, G., & Morehen, J. (2002). Recognition of visually presented musical intervals. *Psychology of Music*, *30*(1), 48–57.

Gobet, F. (1998). Expert memory: A comparison of four theories. *Cognition*, *66*(2), 115–152.

Godoy, R. I., & Jorgensen, H. (2001). *Musical Imagery*. Lisse: Swet and Zeitlinger.

Goebl, W., & Palmer, C. (2006). Anticipatory motion in piano performance. *Journal of the Acoustical Society of America*, *120*, 3002.

Goolsby, T. W. (1994). Eye movement in music reading: Effects of reading ability, notational complexity, and encounters. *Music Perception*, *12*(1), 77–96.

Gordon, E. E. (1993). *Learning sequences in music: Skill, content, and pattern. A music learning theory* (4th ed.). Chicago, IL: GIA.

Guttman, S. E., Gilroy, L. A., & Blake, R. (2005). Hearing what the eyes see: Auditory encoding of visual temporal sequences. *Psychological Science*, *16*(3), 228–235.

Halpern, A., & Bower, G. (1982). Musical expertise and melodic structure in memory for musical notation. *American Journal of Psychology*, *95*(1), 31–50.

Hasegawa, T., Matsuki, K., Ueno, T., Maeda, Y., Matsue, Y., Konishi, Y., et al. (2004). Learned audio-visual cross-modal associations in observed. Piano playing activates the left planum temporale. An fMRI study. *Brain Research, Cognitive Brain Research*, *20*(3), 510–518.

Holcomb, P. J., & Anderson, J. E. (1993). Cross-modal semantic priming: A time-course analysis using event-related brain potentials. *Language and Cognitive Processes*, *8*(4), 379–411.

Jacobsen, O. I. (1942). An analytical study of eye-movements in reading vocal and instrumental music. *Journal of Musicology*, *3*, 197–221.

Just, M. A., Newman, S. D., Keller, T. A., McEleney, A., & Carpenter, P. A. (2004). Imagery in sentence comprehension: An fMRI study. *NeuroImage*, *21*(1), 112–124.

Kennedy, A., Murray, W. S., O'Regan, J. K., & Levy-Schoen, A. (1987). The components of reading time: Eye movement patterns of good and poor readers. In J. K. Regan & A. Levy-Schoen (Eds.), *Eye movements: From physiology to cognition* (pp. 509–520). New York, NY: Elsevier Science.

Kinsler, V., & Carpenter, R. H. S. (1995). Saccadic eye movements while reading music. *Vision Research*, *35*(10), 1447–1458.

Lehmann, A. C., & Ericsson, K. A. (1996). Performance without preparation: Structure and acquisition of expert sight-reading and accompanying performance. *Psychomusicology*, *15*(1–2), 1–29.

Luisa Dematté, M., Sanabria, D., & Spence, C. (2006). Cross-modal associations between odors and colors. *Chemical Senses*, *31*(6), 531–538.

Meister, I. G., Krings, T., Foltys, H., Boroojerdi, B., Muller, M., Topper, R., et al. (2004). Playing piano in the mind – an fMRI study on music imagery and performance in pianists. *Cognitive Brain Research*, *19*(3), 219–228.

Morris, R. K., Rayner, K., & Pollatsek, A. (1990). Eye movement guidance in reading: The role of parafoveal letter and space information. *Journal of Experimental Psychology: Human Perception and Performance*, *16*(2), 268–281.

Parncutt, R., Sloboda, J. A., Clarke, E. F., Raekallio, M., & Desain, P. (1997). An ergonomic model of keyboard fingering for melodic fragments. *Music Perception*, *14*(4), 341–382.

Phillips-Silver, J., & Trainor, L. J. (2007). Hearing what the body feels: Auditory encoding of rhythmic movement. *Cognition*, *105*(3), 533–546.

Rayner, K. (1993). Eye movements in reading: Recent developments. *Psychological Science*, *2*(3), 81–85.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*(3), 372–422.

Rayner, K., & Pollatsek, A. (1997). Eye movements, the eye-hand span, and the perceptual span during sight-reading of music. *Current Directions in Psychological Science*, *6*(2), 49–53.

Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review*, *105*(1), 125–157.

Reisberg, D. (1992). *Auditory imagery*. Hillsdale, NJ: Lawrence Erlbaum.

Robinson, C. W., & Sloutsky, V. M. (2007). Visual processing speed: Effects of auditory input on visual processing. *Developmental Science*, *10*(6), 734–740.

Rounds, B. B., Manley, C. W., & Norris, R. H. (1991). The effect of oculomotor training on reading efficiency. *Journal of the American Optometric Association*, *62*(2), 92–99.

Rugg, M. D., Doyle, M. C., & Melan, C. (1993). An event-related potential study of the effects of within- and across-modality word repetition. *Language & Cognitive Processes*, *8*(4), 357–377.

Schumann, R., & Schumann, C. (2009). *Journal intime*. Conseils aux jeunes musiciens. Paris: Editions Buchet-Chastel. (Original work published 1848)

Shams, L., Kamitani, Y., & Shimojo, S. (2000). What you see is what you hear. *Nature*, *408*, 788.

Shimojo, S., & Shams, L. (2001). Sensory modalities are not separate modalities: Plasticity and interactions. *Current Opinion in Neurobiology*, *11*(4), 505–509.

Sloboda, J. A. (1974). The eye-hand span: An approach to the study of sight-reading. *Psychology of Music*, *2*(2), 4–10.

Sloboda, J. A. (1984). Experimental studies of music reading: A review. *Music Perception*, *2*(2), 222–236.

Sloboda, J. A., Clarke, E. F., Parncutt, R., & Raekallio, M. (1998). Determinants of finger choice in piano sight-reading. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(1), 185–203.

Stewart, L., Henson, R., Kampe, K., Walsh, V., Turner, R., & Frith, U. (2003). Brain changes after learning to read and play music. *NeuroImage*, *20*(1), 71–83.

Thompson, V. A., & Paivio, A. (1994). Memory for pictures and sounds: Independence of auditory and visual codes. *Canadian Journal of Experimental Psychology*, *48*(3), 380–396.

Tinker, M. A. (1946). The study of eye movements in reading. *Psychological Bulletin*, *43*(2), 93–120.

Truitt, F. E., Clifton, C., Jr., Pollatsek, A., & Rayner, K. (1997). The perceptual span and eye-hand span in sight reading music. *Visual Cognition*, *4*(2), 143–161.

Waters, A. J., Townsend, E., & Underwood, G. (1998). Expertise in musical sight reading: A study of pianists. *British Journal of Psychology*, *89*, 123–149.

Waters, A. J., Underwood, G., & Findlay, J. M. (1997). Studying expertise in music reading: Use of a pattern-matching paradigm. *Perception & Psychophysics*, *59(4)*, 477–488.

Weaver, H. E. (1943). A survey of visual processes in reading differently constructed musical selections. *Psychological Monographs*, *55*, 1–30.

Williamon, A., & Egner, T. (2004). Memory structures for encoding and retrieving a piece of music: Aan ERP investigation. *Cognitive Brain Research*, *22*(1), 36–44.

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

Wöllner, C., Halfpenny, E., Ho, S., & Kurosawa, K. (2003). The effects of distracted inner hearing on sight-reading. *Psychology of Music*, *31*(4), 377–389.

Yumoto, M., Matsuda, M., Itoh, K., Uno, A., Karino, S., Saitoh, O., et al. (2005). Auditory imagery mismatch negativity elicited in musicians. *Neuroreport: For Rapid Communication of Neuroscience Research*, *16*(11), 1175–1178.

Zatorre, R. J., Chen, J. L., & Penhune, V. B. (2007). When the brain plays music: Auditory-motor interactions in music perception and production. *Nature Reviews Neuroscience*, *8*, 547–558.

Biographies

Veronique Drai-Zerbib has a PhD in Cognitive Psychology from the Université de Nice Sophia Antipolis, France. She is a researcher at the LASMIC (Université de Nice) and LUTIN (Cité des sciences et de l'industrie, Paris).

Thierry Baccino is Professor in Cognitive Psychology at the Université de Paris VIII, France and is Scientific Director of the LUTIN (Cité des sciences et de l'industrie, Paris).

Emmanuel Bigand is Professor in Cognitive Psychology and is at the Université de Bourgogne, France.

Appendix 1. Listing of excerpts and composers

No. fragment	Extract	Composer
1	Menuette et Aria	Haydn
2	Menuette et Aria	Haydn
3	Menuette et Aria	Haydn
4	Menuette et Aria	Haydn
5	10 pièces - Allegro	Haydn
6	Polonaise	Dussek
7	Vagabond's song 2	Bartok
8	A little song	Kabalevsky
9	Air2	Mozart
10	Andante	Mozart
11	Burleska	Mozart
12	Rondo	Mozart
13	Menuetto	Mozart
14	Rigaudon	Krebs
15	Menuett	Bach

(Continued)

Appendix 1. *(Continued)*

No. fragment	Extract	Composer
16	Sarabande	Mattheson
17	Prelude	Handel
18	The fair	Czerny
19	Une larme	Moussorgsky
20	Sonate n° 11	Cimarosa
21	Six écossaises	Beethoven
22	Vagabond's song	Bartok
23	Minuet	Handel
24	Le moine bourru	Schumann
25	29° Etude	Lemoine
26	10° étude	Lemoine
27	Prélude	Chopin
28	Sarabande	Corelli
29	Fantasia	Telemann
30	Marcia	Mozart
31	Walzer	Brahms
32	Danse	Czerny
33	Le Moulin	Lack
34	Menuet II	Bach
35	Courante	Handel
36	Soldier's march	Schumann
37	Menuette 12	Haydn
38	Menuet	Mozart
39	Air	Mozart