

An fMRI study of music sight-reading

Daniele Schön,^{1,2,CA} Jean Luc Anton,³ Muriel Roth³ and Mireille Besson¹

¹Equipe Langage et Musique, INPC-CNRS, 31 Chemin Joseph Aiguier, I3402 Marseille Cedex 20, France; ²Dipartimento di Psicologia, Università di Trieste, Italy; ³Centre IRMf, CHU Timone, Marseille, France

^{CA}Corresponding Author and Address: danschon@Inf.cnrs.mrs.fr

Received 1 August 2002; accepted 8 October 2002

DOI: 10.1097/01.wnr.0000044224.79663.f5

The brain areas involved in music reading were investigated using fMRI. In order to evaluate the specificity of these areas we compared reading music notation to reading verbal and number notations in a task that required professional pianists to play the notes (in musical and verbal notations) and the numbers displayed on a 5-key keyboard. Overall, the three tasks revealed a similar pattern of activated brain areas. However, direct contrasts between the music notation and the verbal or the numerical notation tasks also revealed specific major foci of activation in the right occipito-temporal junction, superior parietal lobule and the intraparietal sulcus.

We interpret the right occipito-temporal difference as due to differences at the encoding level between notes, words and numbers. This area might be analogous to one described for words, called the visual word form area. The parietal activations are discussed in terms of visuo-motor transcoding pathways that differ for the three types of notations used. Finally, we present a model of music reading that can possibly explain our findings. *NeuroReport* 13:2285–2289 © 2002 Lippincott Williams & Wilkins.

Key words: fMRI; Music; Notational systems; Reading model

INTRODUCTION

Musical notes, in common with words and numbers, can be represented in notational form. Although these notational systems are somewhat different, they all serve the same goal of reading. Whether the verbal and numerical script-dependent reading processes function in similar or different ways has been the matter of debate [1,2]. In the neuropsychological literature, the reported double dissociations, namely selective preservation of Arabic numerals reading together with impairment of letter or word reading [3], and *vice versa* [4], have been taken as evidence for partially independent neural representations.

The main question addressed here is whether the processes involved in music sight-reading are independent from those at play when reading words and numbers. Previous clinical studies mainly report cases of patients with musical disturbances associated with word and/or number disturbances [5–7]. Musicians with alexia for words but not for music have also been described [8–11], and the reverse pattern (preserved language reading with impaired music reading) has been reported recently [12]. Interestingly, Cappelletti *et al.* [12] described a patient who, following a left posterior-temporal lobe lesion and a small right occipito-temporal lesion, showed selective impairments in reading, writing and understanding musical notation, without major disturbances in reading or writing letters, words or numbers.

Surprisingly, most of the literature on music reading comes from cognitive neuropsychology, and very few studies have been carried on using neuroimaging methods.

As the anatomical hypotheses are mostly based on single case studies, often with multiple [12] or large lesions [5,13], uncertainty remains on which areas are necessary for music reading. Another problem is that, with few exceptions, music reading is usually considered as a whole. However, as illustrated in our model of music reading (Fig. 1) three types of transcoding may be involved when a musician reads a score [14,15]: singing-like (i.e. visual to auditory transcoding), playing-like (visual to motor transcoding) or naming-like notes (visual to verbal transcoding). Thus, while some cognitive operations and neural networks might be common to these three types of transcoding, others may well differ. Moreover, the model also illustrates differences between music notation and other notational systems at the encoding level, and different transcoding routes from each type of notation to a given output (e.g. sight-reading).

In the present study we compared music notation to verbal and number notations in a task that required professional pianists to play on a simple 5-key keyboard the notes (either in musical or verbal notations) and the Arabic numbers displayed on a screen (Fig. 2). Compared with neuropsychological single case studies, using fMRI allowed us to test a larger number of normal participants and to localize more precisely the neural networks involved. Moreover, by specifying the task at hand (favouring the visual to motor transcoding) and by comparing reading of music notation with other notational systems, we hoped to shed more light on the specificity of the areas involved in music reading, with respect to previous neuroimaging studies [16,17].

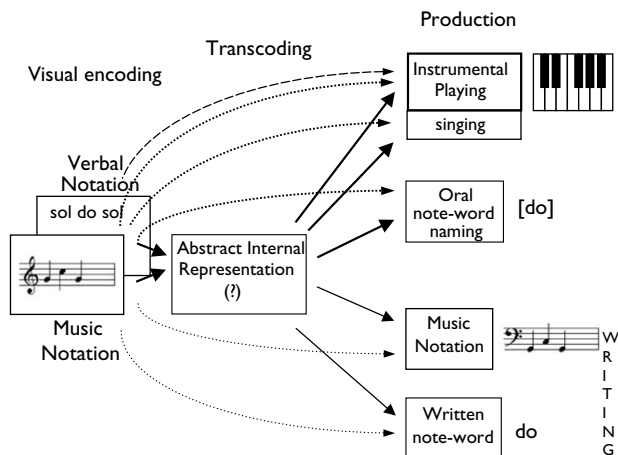


Fig. 1. A minimal model of music reading. Solid lines indicate an indirect route mediated by an abstract internal representation. Dotted lines indicate a direct (asemantic) route from music notation to different types of output. The dashed line indicates that routes different from those associated with music notation, constrain the transcoding from verbal notation. For the sake of simplicity the dashed line is only illustrated for the instrumental playing output.



Fig. 2. Illustration of the mapping between the stimuli presented on the screen in the different experimental conditions and the fingers used for the response. Note that in the experiment, stimuli appeared in random order.

MATERIALS AND METHODS

The design of the experiment comprised three experimental conditions and three control conditions. In all conditions stimuli were visually presented one at a time, in a pseudo-random order, for 800 ms with an inter-stimulus interval of 150 ms. In condition 1, the stimuli were in music notation (five notes, from do to sol, the thumb corresponding to do). In condition 2, the stimuli were in verbal notation (five notes, from do to sol). In condition three, the stimuli were in Arabic number notation (5 numbers, from 1 to 5). In the experimental conditions the task was to play with the right hand on a 5-key keyboard the stimuli successively displayed. In the control conditions, subjects had to press a button (with the fourth finger) each time a stimulus

appeared. Note that control conditions do not control for motor activation, as it is known that sequential and repetitive movements produce qualitatively different activations. However, the important point is that the finger movements are identical in the experimental conditions and that, consequently, the motor activation should be similar. The control conditions were mainly aimed at controlling for the differences in the visual appearance of the stimuli. Thus, the stimuli in the control conditions visually matched those in the experimental conditions: a quarter-pause on the musical staff, a short word (chut, meaning silence), and a zero. Visual stimulation was synchronized with fMRI acquisition. Each block comprised 26 stimuli (pseudo-randomized in the experimental conditions). For each condition, 10 blocks were run, in a pseudo-random order, over three scanning sessions. In order to reduce task-switching-related activation at the beginning of each block, instructions were first displayed (3500 ms) to indicate which type of notation will be presented. The instructions were treated apart in the statistical model.

Imaging was performed using a 3T whole-body imager MEDSPEC 30/80 ADVANCE (Bruker). High-resolution structural T1-weighted images were acquired for all participants to allow precise anatomical localisation ($1 \times 0.75 \times 1.22$ mm). The anatomical slices covered the whole brain and were acquired parallel to the anterior-posterior commissure (AC-PC) plane. The functional images were acquired using a T2*-weighted echo-planar sequence at 26 axial slices (repetition time 2.2 s, interleaved acquisition, slice thickness 4 mm, inter-slice gap 1 mm, 64×64 matrix of 3×3 mm voxels). The slices were parallel to the AC-PC plane, and covered the whole brain. For each session, the scanner was in the acquisition mode for 10 s before the experiment began, to achieve steady-state transverse magnetisation.

Statistical parametric mapping software (SPM99) [18] was used for image processing and analysis. The functional images were interpolated in time to correct phase advance during volume acquisition, and realigned to the first image of the first session. In order to compute multi-subject analysis, the anatomical references and the realigned functional images of all subjects were transformed (non-linear transformations) into a common standard space using the Montreal Neurological Institute template. The functional data were then spatially smoothed (3D Gaussian kernel: $9 \times 9 \times 9$ mm) and temporally filtered, using a 120 s period high-pass filter and a Gaussian low-pass filter with a 4 mm of full width at half maximum (FWHM). A general linear fixed-effect model was applied to the time course of the functional signal at each voxel. Each condition for each subject was modelled by one reference waveform (boxcar convolved with a canonical hemodynamic response function).

Results of the conjunction analyses between subjects, in a fixed effect model [19] are reported below, using a significance threshold for active voxel of $p = 0.05$ (corrected, FDR [20]). Experiments were performed on nine healthy, right-handed volunteers (four women, five men) aged 24–50 years, all with minimum 12 years of piano playing experience. All subjects gave informed consent to the experimental procedure, as required by the Helsinki declaration.

RESULTS

As expected, T contrasts between each playing condition and its own control showed that non-specific visual processing common to the control and the experimental conditions was subtracted, so that no residual signal was left in the primary visual areas. Overall, a similar pattern of brain areas is activated by the three notational systems (Fig. 3), namely the parietal lobes bilaterally (including the superior parietal gyrus, the angular gyrus and the supra-marginal gyrus), the sensorimotor cortex of the left hemisphere (contralateral to the hand used to play), and the right cerebellum. Direct contrasts were computed between the music notation and the verbal or the numerical notations, exclusively masked ($p = 0.01$) by the signal within the music

control condition, so as to identify the brain areas with larger signal in the music sight-reading task. Results showed two major foci of activation for both music *vs* words and music *vs* numbers: one in the right superior parietal lobule (SPL) and the other in the intraparietal sulcus (IPS), mesial to the supramarginal gyrus (Fig. 4). Another minor focus of activation for both contrasts was also found in the right visual cortex, close to the occipito-temporal junction.

DISCUSSION

Previous work by Sergent *et al.* [16] showed bilateral activations of the extrastriate visual areas (areas 18) and a

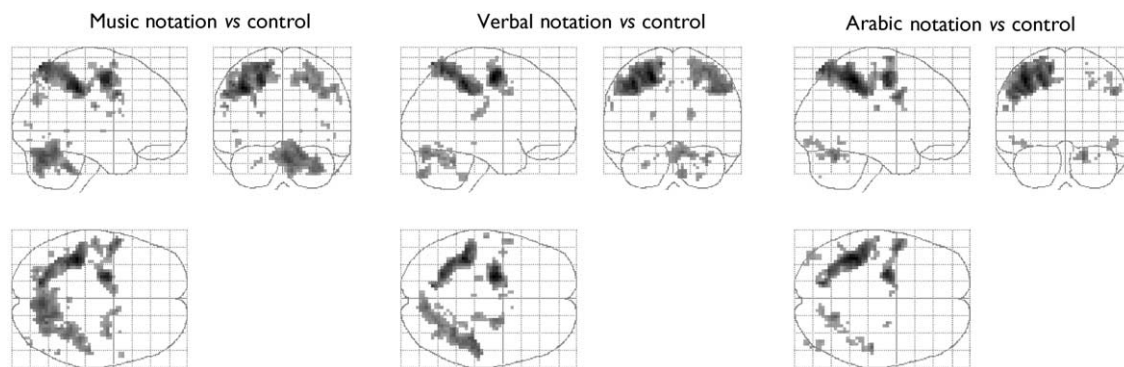


Fig. 3. Comparison of each playing task vs its own control condition ($p < 0.05$).

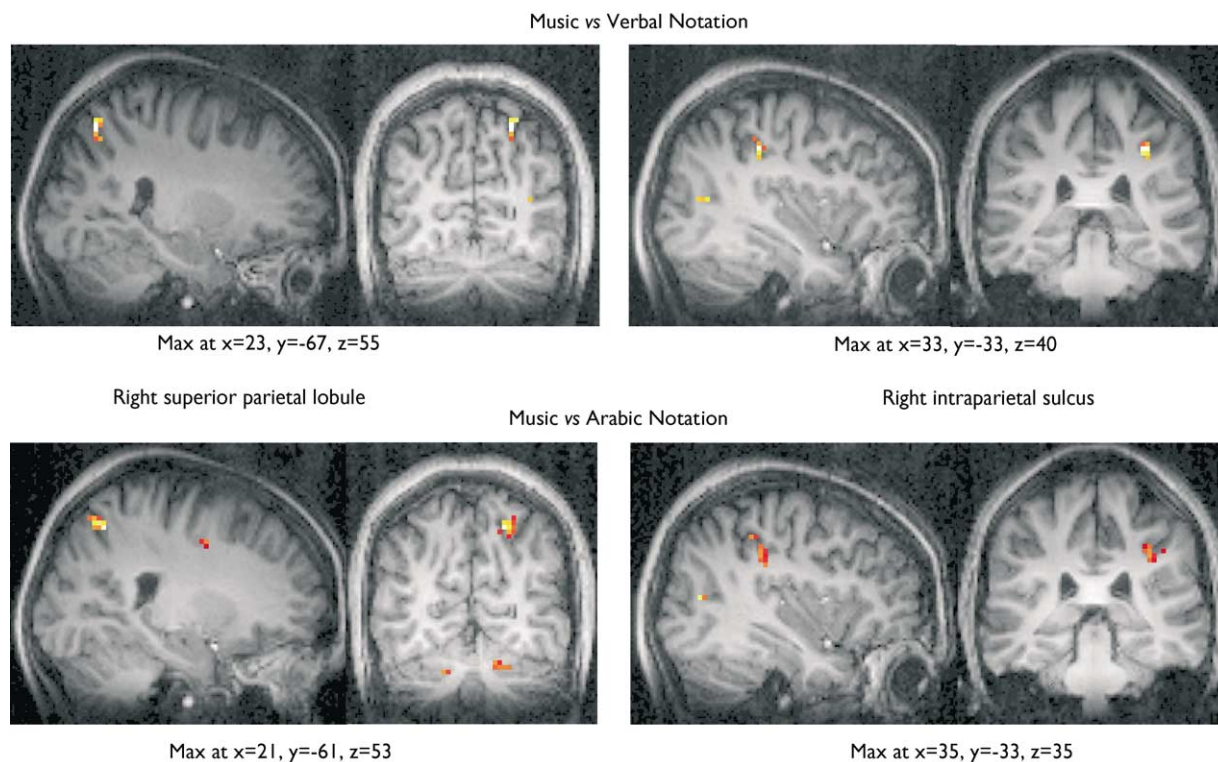


Fig. 4. Areas significantly more active ($p < 0.05$) while reading music notation relative to verbal and number notations. Both contrasts are exclusively masked by music control. MNI coordinates converted to Talairach coordinates (<http://www.mrc-cbu.cam.ac.uk/imaging/>) are shown for the voxel with the highest signal for each contrast in the two parietal sites.

left occipito-parietal activation when musicians were reading a score. However, since the control stimuli, visual dots, were not visually matching the musical score, no strong claim can be made regarding the specificity of these visual areas for music sight-reading. Moreover, the reading task was not clearly defined. Insofar as musicians were not playing the score, they might have used one or several of the previously described ways of reading music (Fig. 1). By contrast, the visual control stimuli in our study were closely matching the stimuli in the experimental conditions, and consequently, no extrastriate visual areas were found activated when contrasting the music control and experimental conditions. Only a small focus was found at the right occipito-temporal junction, when contrasting music reading with its control, even more evident when contrasting music with words and number notation. It is interesting to note that this same focus was found by Nakada *et al.* [17]. These authors compared the activation pattern associated with music score reading with that associated with language reading (English and Japanese). An area within the right occipital cortex (adjacent to the occipital sulcus) was identified as being specifically activated by reading music scores. However, the reading task used by the authors was again not clearly specified, and we are confronted with the same interpretative problem as mentioned for the Sergent study [16]. Nonetheless, this region of the right occipital cortex seems to be important since Cappelletti's *et al.* [12] patient, who was completely unable to read music, also had a small right occipito-temporal lesion.

The contrasts between music notation and verbal or number notations again revealed a right occipito-temporal activation. The most likely interpretation of this difference is that, in music, the pitch of the notes is coded by their position, while letters and numbers are coded according to their form. Moreover, by contrast with words and numbers, each note is coded with respect to its position on a meaningful background, the staff, and might be read in relation one to the other. Thus, with respect to our model of music reading, this right occipital difference would be due to differences at the encoding level between notes, words and numbers (Fig. 1). This area might be the musical functional homologue of the visual word form area involved in prelexical encoding of written words [21,22], and located in the middle portion of the left fusiform gyrus.

Most importantly for the aim of the present study is the finding of a differential parietal activation. In the study by Sergent *et al.* [16], the authors planned a condition very similar to the one used in our experiment. They asked participants to sight-read, play and listen all together. Then, they contrasted this condition with one where participants were reading a score and listening (without playing). They found a bilateral activation of the superior parietal lobules (area 7). They argue that these areas of the parietal cortex are strategically placed to mediate the sensorimotor transformations for visually guided skilled actions and finger positioning. The lack of a condition with a sensorimotor transformation of a different type prevented concluding whether these areas are general purpose or partly specific to music playing. Hence the comparison of different notational systems in the present study. One may argue that, even if the subjects are performing the same task (i.e. attributing a motor response to each presented visual stimulus), the rules that are associated with and constrain the information

represented by the note on the staff are different from those for numbers and words. Indeed, there is some evidence that, within the parietal lobe, verbal and non-verbal (numerical) processes involve areas that are partially non-overlapping [23]. Even within music these transformation rules might be subjected to fine changes, such as when a musician has to change from bass clef to treble clef [14]. Thus, according to our model, the residual signal found in music, once subtracted that due to verbal notes or numbers playing, can be reasonably explained by the different rules that are at work (see Fig. 1, dotted *vs* dashed line). The present results also speak to the issue of brain plasticity in professional musicians [24]. Reading from music notation is certainly more commonly used by musicians than reading from verbal or numerical notations. It is thus possible that more specific neural networks support visuomotor transformation when music notation is used than when verbal or number notations are used. Interestingly, the IPS has been found to mediate the processing of sensorimotor integration of precisely tuned finger movements in humans [25] and to control for the endogenous allocation and maintenance of visuospatial attention [26]. It is, therefore, not surprising that this area was also strongly involved in music sight-reading. Finally, note that the right-sided lateralization of the foci described in the occipito-temporal and parietal cortex might be linked to the right lateralization often described for auditory music processing [27]. However, such speculations should be considered with caution and further studies are necessary to find the link(s) between music notation and the complex auditory perception of a music masterpiece.

CONCLUSION

This study shows that, when playing from music notation, well-defined right parietal regions are more involved than when transcoding from verbal or number notations. However, more research is needed to disentangle the precise role of the two parietal foci found in the present study with respect to the model proposed. Different ways of reading music coexist and can eventually be at work at the same time, and we do not yet know whether the same cerebral parietal networks would also be involved when reading musical notation in order to sing or name the notes. Moreover, music reading, as proposed for number transcoding, may involve a semantic transcoding route through abstract internal representations [28], a direct route through asemantic transcoding algorithms [1], or both [29,30]. Further experiments will be specifically designed to address these issues. Finally, we should also keep in mind that musicians are highly trained in music notation reading. A dedicated neural network might be responsible for this sophisticated skill that allows transforming music notation into a precise motor response.

REFERENCES

1. Deloche G and Seron X. Numerical transcoding: a general production model. In: Deloche G and Seron X. *Mathematical disabilities: a cognitive neuropsychological perspective*. Hillsdale, NJ: Lawrence Erlbaum; 1987.
2. Cohen L, Dehaene S, Chochon F *et al.* *Neuropsychologia* **38**, 1426-1440 (2000).
3. Anderson SW, Damasio AR and Damasio H. *Brain* **11**, 749-766 (1990).

4. Cipolotti L. *Cogn Neuropsychol* **12**, 313–342 (1995).
5. Fasanaro AM, Spitaleri DL and Valiani R. *Music Percept* **7**, 259–272 (1990).
6. Horikoshi T, Asari Y, Watanabe A *et al.* *Cortex* **33**, 187–94 (1997).
7. Kawamura M, Midorikawa A and Kezuka M. *Neuroreport* **11**, 3299–3303 (2000).
8. Assal G and Buttet J. *Rev Neurol* **139**, 569–574 (1983).
9. Basso A and Capitani E. *Neurol Neurosurg Psychiatry* **48**, 407–412 (1985).
10. Signoret JL, Van Eeckhout P, Poncet M *et al.* *Rev Neurol* **143**, 172–181 (1987).
11. Brust JC. *Brain* **103**, 367–92 (1980).
12. Cappelletti M, Waley-Cohen H, Butterworth B *et al.* *Neurocase* **6**, 321–332 (2000).
13. Stanzione M, Grossi D and Roberto L. *Music Percept* **7**, 273–284 (1990).
14. Schön D, Semenza C and Denes G. *Cortex* **37**, 407–421 (2001).
15. Schön D and Besson M. *Neuropsychologia* **40**, 868–878 (2002).
16. Sergent J, Zuch E, Terriaj S *et al.* *Science* **257**, 106–109 (1992).
17. Nakada T, Fujii Y, Suzuki K *et al.* *Neuroreport* **9**, 3853–3856 (1998).
18. Friston KJ, Holmes AP, Poline JB *et al.* *Neuroimage* **2**, 45–53 (1995).
19. Friston KJ, Holmes AP, Price CJ *et al.* *Neuroimage* **10**, 385–396 (1999).
20. Genovese CR, Lazar NA and Nichols TE. *Neuroimage* **15**, 772–786 (2002).
21. Cohen L, Lehericy S, Chochon F *et al.* *Brain* **125**, 1054–1069 (2002).
22. Dehaene S, Le Clec'H G, Poline JB *et al.* *Neuroreport* **13**, 321–325 (2002).
23. Cohen L, Dehaene S, Naccache L *et al.* *Brain* **123**, 291–307 (2000).
24. Munte TF, Altenmuller E and Jancke L. *Nature Rev Neurosci* **3**, 473–478 (2002).
25. Binkofski F, Dohle C, Posse S *et al.* *Neurology* **50**, 1253–1259 (1998).
26. Corbetta M, Kincade JM and Shulman GL. *J Cogn Neurosci* **14**, 508–523 (2002).
27. Zatorre RJ, Belin P and Penhune VB. *Trends Cogn Sci* **6**, 37–46 (2002).
28. McCloskey M, Caramazza A and Basili A. *Brain Cogn* **4**, 171–196 (1985).
29. Dehaene S and Cohen L. *Math Cogn* **1**, 83–120 (1995).
30. Cipolotti L and Butterworth B. *J Exp Psy Gen* **124**, 375–390 (1995).

Acknowledgements: This research was supported by a grant from the International Foundation for Music Research to M.B. (IFRM: RA #194). D. S. was supported by the IFMR to conduct this research (2001–2002).