INTERNATIONAL JOURNAL OF PSYCHOLOGY, 2000, *35* (6), 270–278

Divided attention in music

E. Bigand
Université de Bourgogne,
Dijon, France

S. McAdams
Université René Descartes,
and Institut de Recherche et de
Coordination Acoustique/Musique
(IRCAM-CNRS), Paris, France

S. Forêt
Université de Bourgogne, Dijon, France

Two models have been advanced to account for the apparent ease with which attention can be divided in music: a "divided attention" model postulates that listeners effectively manage to follow two or more melodic lines played simultaneously. According to a "figure-ground model," the harmonic coherence of Western polyphonies allows a listener to focus on one melody while staying aware of the other melody, which acts as a background. This figure-ground processing compensates for the inability to divide attention. The present study was designed to further investigate these two models. Participants were required to detect melodic errors in two familiar nursery tunes played simultaneously an octave apart. The divided-attention model predicted that this task would be easily performed by participants, irrespective of the key of the nursery tune. The figure-ground model predicted better performance when the keys of the tunes were identical or closely related. None of these predictions was fully supported by the data, leading us to propose a new "integrative model" of listening to polyphonic music.

Deux modèles ont été proposés pour rendre compte de l'apparente facilité à partager son attention lors de l'écoute musicale: selon un modèle "d'attention partagée", l'auditeur parviendrait à suivre deux ou plusieurs lignes mélodiques jouées simultanément. Selon un modèle "figure-fond", la cohérence harmonique des polyphonies occidentales permettrait à l'auditeur de focaliser son attention sur une mélodie tout en restant attentif à l'autre mélodie qui agirait comme un fond harmonieux. Ce processus d'organisation de la polyphonie en figure-fond compenserait l'incapacité à diviser l'attention. L'étude présente fut conçue pour approfondir ces deux modèles. Les sujets devaient détecter des erreurs mélodiques dans deux mélodies populaires jouées simultanément à deux octaves différentes. Selon le modèle "d'attention partagée", cette tâche devait être facilement réalisée, quelque soit la tonalité respective des mélodies. Le modèle "figure-fond" prédisait de meilleures performances lorsque les tonalités des mélodies étaient identiques ou proches. Aucune de ces prédictions n'a été pleinement confirmée par les données ce qui conduit à développer un nouveau modèle "d'écoute intégrative" de la polyphonie.

Human beings can usually attend to only one message at a time. For example, it is hard to listen to one person while simultaneously conversing with another one. Subjects required in laboratory tasks to track two linguistic messages at the same time usually manage to focus on one spoken message but do not process the other (Cherry, 1953). Dividing one's attention raises difficulties not only for spoken language, but also for a number of other types of temporal information structures. For example, Neisser and Becklen (1975) reported that participants performed well when playing one video game while ignoring a second one superimposed on the screen (selective attention task). However, performance severely dropped when participants were required to play both video games at the same time (divided attention task).

It has been initially suggested that human attention is limited to a single channel at any one time (Broadbent, 1958). A less extreme view is that the dividing of attention is less difficult when the simultaneous messages to be followed tap into different cognitive processes. For example, Hirst and Kalmar (1987) required participants to detect errors in simultaneous messages sent to each ear separately (dichotic listening). Messages to be processed were a set of letters forming a word or a nonword (A-M-E-R-I-C-A-N versus A-M-R-I-D-A-N) or a regular or irregular set of numbers (2-4-6-8-10-12 versus 2-4-6-9-10-12). In one condition, simultaneous messages were similar (two different sets of letters, for example), in the other simultaneous messages were different (one set of letters and one set of numbers). Participants detected only a few errors in the former condition but had better performance in the latter one. Training has also been shown to be an important factor in the ability to divide one's attention. It is obvious, for example, that new drivers

---

Requests for reprints should be addressed to E. Bigand, Laboratoire des Apprentissages et du Développement (CNRS). Bld Gabriel, Université de Bourgogne, 21000, Dijon, France (E-mail: bigand@satie.u-bourgogne.fr).

http://www.tandf.co.uk/journals/pp/00207594.html

have greater difficulty in driving and speaking at the same time. Performing both tasks simultaneously becomes easy after a while. Spelke, Hirst, and Neisser (1976) trained students to read a text while writing words pronounced by the experimenter. At the beginning of the training session, their performance in both tasks was very low (i.e. slow reading tempo and very poor transcription of the dictated words). After 6 weeks of training, they managed to read as quickly as they did without a second task, and the writing performance was much better.

Music is a thought-provoking domain for research on divided attention. Indeed, most of the music we listen to is made up of simultaneous, similar sound streams (i.e., multivoiced or polyphonic music). Serious vocal music relies on four different voices referred to as soprano, alto, tenor, and bass voices, and instrumental pieces often contain more than 10 different voices (as in symphonies, concerti, or chamber music). Multivoiced music occurred very early in the history of Western music and is present in numerous folk traditions around the world. Even nursery tunes are often played simultaneously with a second melody, embellishing the tune. If music presented the same difficulty for divided attention as has been observed in other domains (spoken language, video games, and so on), polyphonic music would probably not have become as developed as it has. Is music, then, a specific domain of cognition in which attentional processes can be easily divided? Does it present structural characteristics that make it possible to compensate for the attentional constraints observed in several other domains? If so, what are these characteristics? The purpose of the present study was to investigate the dividing of attention between two voices in polyphony.

Perceiving polyphony first requires one to segregate sounds into different streams. Several factors intervene at this stage of processing (see Bregman, 1993, for a review, and Bregman, 1990, for an exhaustive account). The frequency distance between the sounds of each stream is the most important factor if the sounds have the same spectral content (Miller & Heise, 1950; van Noorden, 1975). Played in the same pitch range, the two voices of a polyphony played by the same instrument are not easily segregated: the melodies making up the polyphony would thus be difficult to track because they are fused into a single auditory stream (see Figure 1). Segregation is a necessary condition for perceiving multivoiced music. Once segregation has been achieved, however, perceiving polyphonic music raises the divided attention problem. For example, Dowling (1973) reported that once the pitch range of two interleaved familiar tunes are separated enough (e.g., by one octave or more), participants perceived them as separate streams, but had difficulty identifying both of them.

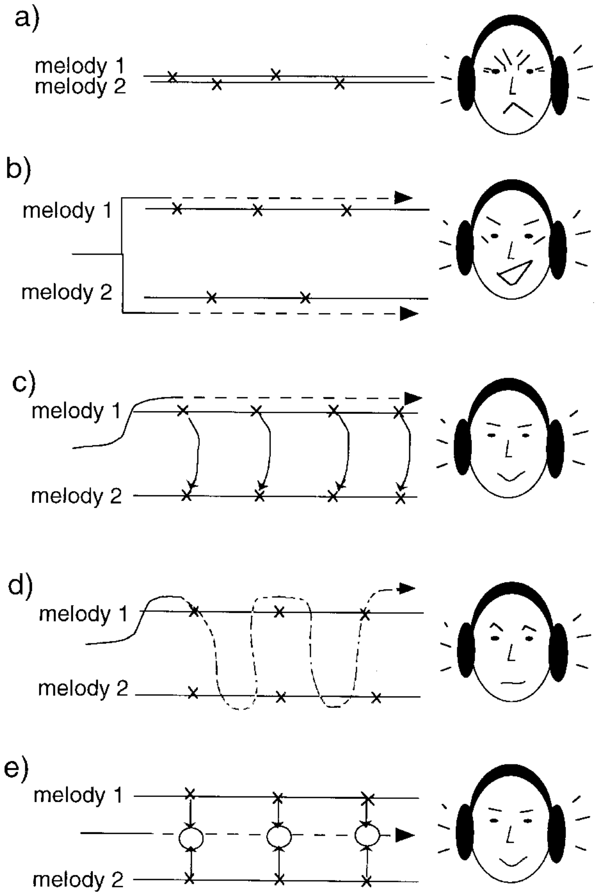

Figure 1. Different perceptual processes involved in listening to polyphony: (a) obligatory integration, (b) divided attention, (c) figure-ground, (d) attentional switching, (e) voluntary integration.

Several studies have been devoted to the perception of multivoiced music (Huron, 1989; Huron & Fantini, 1989; Palmer & Holleran, 1994; Thompson, 1993), but only a few have directly investigated divided attention in polyphonic listening. In an early study, Kahneman (1973) observed that listeners had great difficulty following two songs at once. In this experiment, participants had to track one tune (by singing its tones) presented in one ear while ignoring the other. At the end, they were asked about the identity of the ignored tune. None managed to answer correctly, suggesting that music encounters divided attention difficulties similar to those of spoken language. More recent research, however, has provided evidence that a divided attending task is rather well performed with music. Gregory (1990) presented short extracts of polyphonic music to participants who were required to decide whether or not a subsequent melody was present in the polyphonic extract. Recognition accuracy was high (up to 91%), suggesting that the different melodic lines of the polyphony can be perceived simultaneously. Moreover, no difference in recognition accuracy was observed for polyphonies with two or three melodic lines. A critical finding was that recognition accuracy depends on several factors such as key relatedness, timbre, pitch range, and tempo. Recognition was notably more accurate when the melodies of the polyphony were closely related in key, although the size of the difference was not very large (75% v. 72%).

An effect of key relatedness on divided attention in music was also reported by Sloboda and Edworthy (1981). In their experiment, participants had to localize in a musical score one melodic error occurring in a well-learned, two-voice polyphony. The error detection task was chosen as a way of monitoring the degree of attention allocated to the two simultaneous melodies: If the participants followed both melodies simultaneously, they would detect most of the errors, but if they used other perceptual processes (such as quickly switching their attention between the different tunes), they would detect only half of the errors. When both melodies of the polyphony were played in the same key, correct location of the melodic error was moderately high (69%). The number of correct responses decreased (56%) when the melodies were played in different but related keys (C and G major) and were somewhat low (47%) when they were played in unrelated keys (C and F# major). When the task simply consisted of deciding whether or not a melodic error occurred in the polyphony (Expt. 2), percentages of correct responses rose to 81% for musicians and 80% for nonmusicians in the same key condition. In the worst case (i.e. when the voices were in unrelated keys), percentages of correct responses remained moderately high (72% and 64% for musicians and nonmusicians, respectively).

These previous studies by Gregory (1990) and Sloboda and Edworthy (1981) demonstrate that a divided attention task can be successfully performed in music by comparison to what is usually observed in other domains. Two kinds of explanation may be advanced to account for this specificity of music (Figure 1b and 1c). First, it may be postulated that participants effectively managed to follow two or more messages played simultaneously in music ("divided attention" model). Music may thus be a specific domain of cognition (Peretz & Morais, 1989) relying on attentional processes that do not share the same constraints as those encountered in other domains.

An alternative view is to say that our attention is limited to a single channel in music as in other domains, but that the specificity of music is in its employment of perceptual mechanisms that compensate for the inability to divide attention. Polyphonic music exhibits strong differences with simultaneous sequences of environmental sound events. Melodic lines of a polyphony are conceived by composers to create a coherent overall pattern. The musical rules of counterpoint not only assure a perceptual coherence between the notes belonging to the same stream (horizontal coherence) but also constrain the harmonic relationships between streams (vertical organization).

The fact that divided attention gets easier when the melodies of a polyphony are in the same key suggests that perceptual mechanisms could benefit from this harmonic coherence. According to Sloboda and Edworthy (1981), harmonic coherence might allow the successive processing of each melody as a figure detached from a harmonic background. As in Rubin's "face-vase" drawings, this harmonic coherence allows a listener to focus on one melody while staying aware of the other melody which acts as a background ("figure-ground" model). In other words, one melodic line is focused on, with the tone of the other acting as part of the vertical harmony of the polyphony (Figure 1c). This figure-ground processing of the polyphony makes it possible to detect a melodic error even if it occurs in the nonfocused melody, since the error would then modify the overall figure-ground relation. When the melodies are played in two different (more or less distant) keys there is not enough harmonic coherence between the voices to enable processing of one melody as a background melody of the other one. In such a situation, the sole possibility to track the two melodies at the same time would be to switch back and forth between them (Figure 1d).

The Sloboda and Edworthy (1981) study was not designed to disentangle a true divided attention model from a figure-ground processing model. Given that there was only one melodic error to detect per trial and that this error created a dissonant interval most of the time, the divided attention task may have been partially confounded in their experiment with a dissonant interval detection task. Since the number of dissonant intervals necessarily increased with the key distance between the voices (i.e. harmonically unrelated keys have fewer common notes than related keys), localizing the melodic error may have been more difficult in the unrelated key condition because of this confound. The present study was designed to further investigate the two models using an experimental method that differs from that of Sloboda and Edworthy in several ways. In the present study, participants were presented with familiar nursery tunes played simultaneously an octave apart. Each tune was well known (Bigand, 1990), but participants had never heard their polyphonic mixture prior to the experiment. In order to force participants to follow each tune simultaneously, two melodic errors were introduced in each of the tunes. These four melodic errors never created dissonant intervals when the tunes were played in same- and related-key conditions. In addition, an on-line procedure was used that made it possible to register hits and false alarms in real time. Finally, in order to assess the potential influence of musical training on the dividing of attention between voices, the performance of musically naive participants was compared to that of music students (Experiment 1). According to a true divided attention model, participants would have no difficulty detecting most of the melodic errors contained in both upper and lower tunes. In contrast, a figure-ground model predicts moderate performance since the melodic errors never introduce dissonant intervals, with better performance in the same-key condition than in the closely related (near-key) and unrelated (far-key) conditions. Neither of these two models makes specific predictions about false alarm responses.

EXPERIMENT 1

Participants

Twenty-four participants, with ages varying between 19 and 23 years, took part in the experiment. Half of them had received at least 10 years of intensive training in music (i.e., music theory, ear training, and instrumental performance) and will be referred to hereafter as "musicians." The other half had never had any formal instrumental or theoretical training and will be referred to as "nonmusicians." Nonmusicians were familiar with music and reported listening to music about 1 hour per day on average.

Material

Four equally well-known French nursery rhymes were chosen for this experiment (Bigand, 1990, Expt. 4). They were organized into two pairs ("*Gentille alouette—J'ai du bon tabac*"; "*Sur le pont d'Avignon—A la claire fontaine*") so that the two tunes of a given pair would be equally recognizable when played at the same tempo (i.e., 50 quarter notes per minute for the first pair; 70 quarter notes per minute for the second pair). In addition, these nursery tunes were paired so that they would not create a large number of dissonant intervals when played in the same key. To avoid fusion, the melodies of one pair were played simultaneously at a mean pitch difference of about an octave (Figure 2). The pitch ranges of the melodies (high/low) were counterbalanced across participants.

Three key conditions were defined that did not introduce confounded changes in pitch range. In the same-key condition, melodies of the pair were played in the G major key. In the near-key condition, the upper melody was played in G major and the lower in A major. Two steps separate the keys of G and A major on the circle of fifths that represents key distance (see Krumhansl, 1990, for a psychological account of key distance). In the far-key condition, the upper melody was played in G and the lower in A-flat. Five steps separate these keys on the circle of fifths. Two errors were placed in each melody at different moments (see Figure 2). These errors never created dissonant intervals when the melodies were in the same- and related-key conditions.

All the melodies were played with sampled piano sounds produced by the EMT10 Yamaha Sound Expander. The Yamaha sampler was controlled through a MIDI interface by a Macintosh computer running Performer software. Velocity (a parameter related to the force with which a key is struck) was held strictly constant for all tones. Participants were allowed to adjust the output of the amplifier to a comfortable listening level. There was no silence between the offset of a tone and the onset of the succeeding one.

Procedure

In the first part of the experiment, participants were presented with a correct version of each of the tunes played in isolation. This first part was supposed to
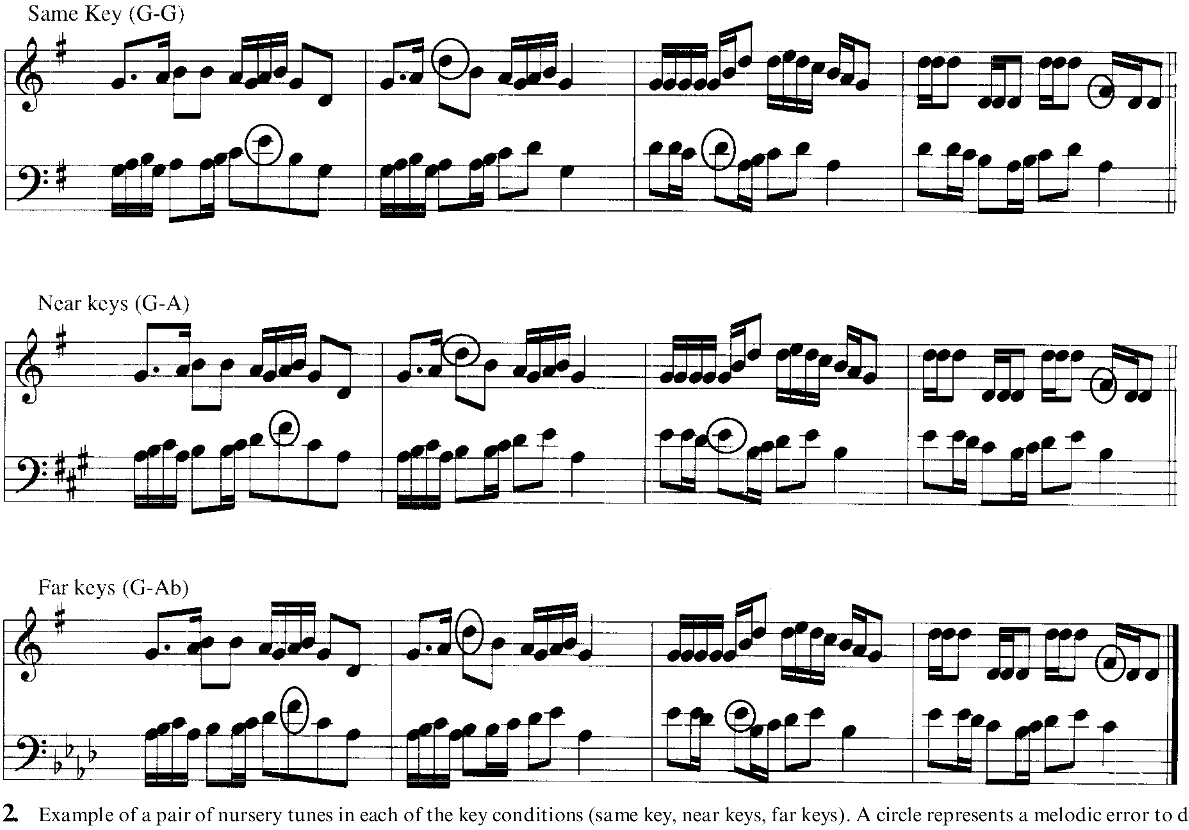
Figure 2. Example of a pair of nursery tunes in each of the key conditions (same key, near keys, far keys). A circle represents a melodic error to detect.

refresh their memory for these tunes and it was verified that all participants were able to name the four melodies prior to the experiment. Participants were then informed that they would listen to two tunes played simultaneously that contained four melodic errors. Their task consisted of pressing a key on an electronic piano keyboard as soon as they detected a melodic error in either the upper or the lower tune (divided attention task). The keyboard was connected to the computer that recorded the participants' responses with 10 ms accuracy. Participants were first trained on the task with a pair of French nursery tunes ("Au clair de la lune"/"Il était un petit navire") that was presented twice for each of the three key conditions. Then, in separate blocks, the two experimental pairs of melodies were used. Each pair was presented in all three key-distance conditions in two successive runs. The order of presentation of key-distance conditions within a run and of melody pairs across runs was counterbalanced across participants. The last part of the experiment consisted of presenting each tune separately with the melodic errors. The participant's task was still to detect these errors as quickly as possible. This last control condition was designed to test whether the melodic errors introduced in the tune were detectable when the tunes were played in isolation.

Results

Recording participants' responses with the computer allows each key press to be localized in the musical score. Since the responses were made in real time, a predetermined time-window equal to the duration of an eighth-note was used to discriminate between hits and false alarms. Key presses were systematically considered as hits when they occurred within the time-window starting with the onset of the tone with the pitch error (Figure 3). They were considered as false alarms when they occurred outside the time-window. This criterion was quite constraining and may have rejected some hits resulting from responses made during the execution of the next note in the polyphony. However, the analysis of the temporal distribution of false alarms indicates that they were no more frequent immediately following a melodic error than elsewhere in the polyphony. Less than 5% of false alarms were located in the time intervals that immediately follow the melodic errors, suggesting that there was no trade-off between speed and accuracy. In addition, participants managed to detect most of the melodic errors according to this criterion when the tunes were played separately (98% and 97% of the errors for nonmusicians and musicians, respectively).

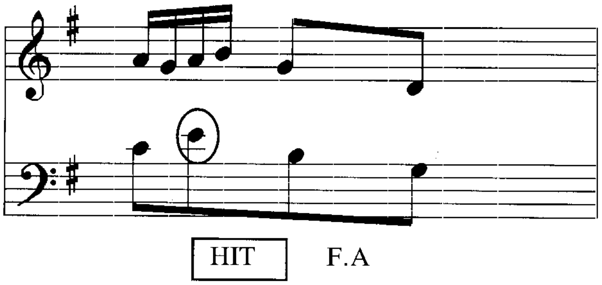

Figure 3. Time-window used to discriminate between hits and false alarms: Participants' responses were considered as hits when they occurred within the time-window starting with the onset of the tone with the pitch error (circle). They were considered as false alarms when they occurred outside the time-window.

Figure 4 (top) represents the average number of hits out of four possible as a function of musical expertise, run, and key conditions. Musicians detected on average 2.08 errors (52%) and nonmusicians 1.48 errors (37%). The number of hits did not vary as a function of the key condition nor as a function of the run number. A 2 × 2 × 3 (Musical expertise × Run × Key condition) ANOVA, performed with the first factor as the between-subject variable, revealed only a significant effect of musical expertise, $F(1, 22) = 7.27$, $p < .02$. There was no significant effect of the run nor key condition. The number of hits slightly increased for musicians from run 1 (1.967) to run 2 (2.193), but this difference was not significant, $F < 1$. A further analysis indicated that musicians detected on average 1.09 errors in the upper melody and 0.98 errors in the lower melody. In contrast, nonmusicians detected 1.07 errors in the upper voice, but only 0.41 errors in the lower one. This suggests that this latter group mainly focused on the most salient melody, thus simplifying the divided attention task into one of selective attending.

Figure 4 (bottom) represents the average number of false alarms as a function of musical expertise, run, and key condition. A 2 × 2 × 3 (Musical expertise × Run × Key condition) ANOVA, performed with the first factor as the between-subject variable, revealed a significant effect of musical expertise, $F(1, 22) = 11.48$. $p < .01$, with numerous false alarms for musicians. There was a main effect of key condition, $F(2, 44) = 12.99$, $p < .01$: The number of false alarms increased as a function of the distance between the keys. There was a significant Key condition × Musical expertise interaction with a significant effect of key condition in musicians, $F(2, 44) = 8.86$, $p < .001$. Contrast analysis indicated no significant effect of key condition for nonmusicians, $F < 1$. In addition, the false alarm rate decreased during the second run, $F(1, 22) = 5.99$, $p < .05$. This effect of run was mostly observed with musicians, $F(1, 22) = 5.39$, $p < .05$. The number of false alarms slightly decreased in nonmusicians from run 1 (0.387) to run 2 (0.303), but this difference was not significant.

Dissonant intervals are more numerous when the distance between the keys of the melodies increases. For the present stimuli, there were eight dissonant intervals over the two pairs of melodies in the same-key condition. This number increases to 16 in the near-key condition and to 32 in the far-key condition. Therefore, it is likely that the effect of key condition on the number of false alarms was caused by the presence of an increasing number of dissonant intervals. To examine this interpretation further, we computed the percentage of false alarms observed for all of the pitch intervals of the two polyphonies. As can be seen in Table 1, most of the false

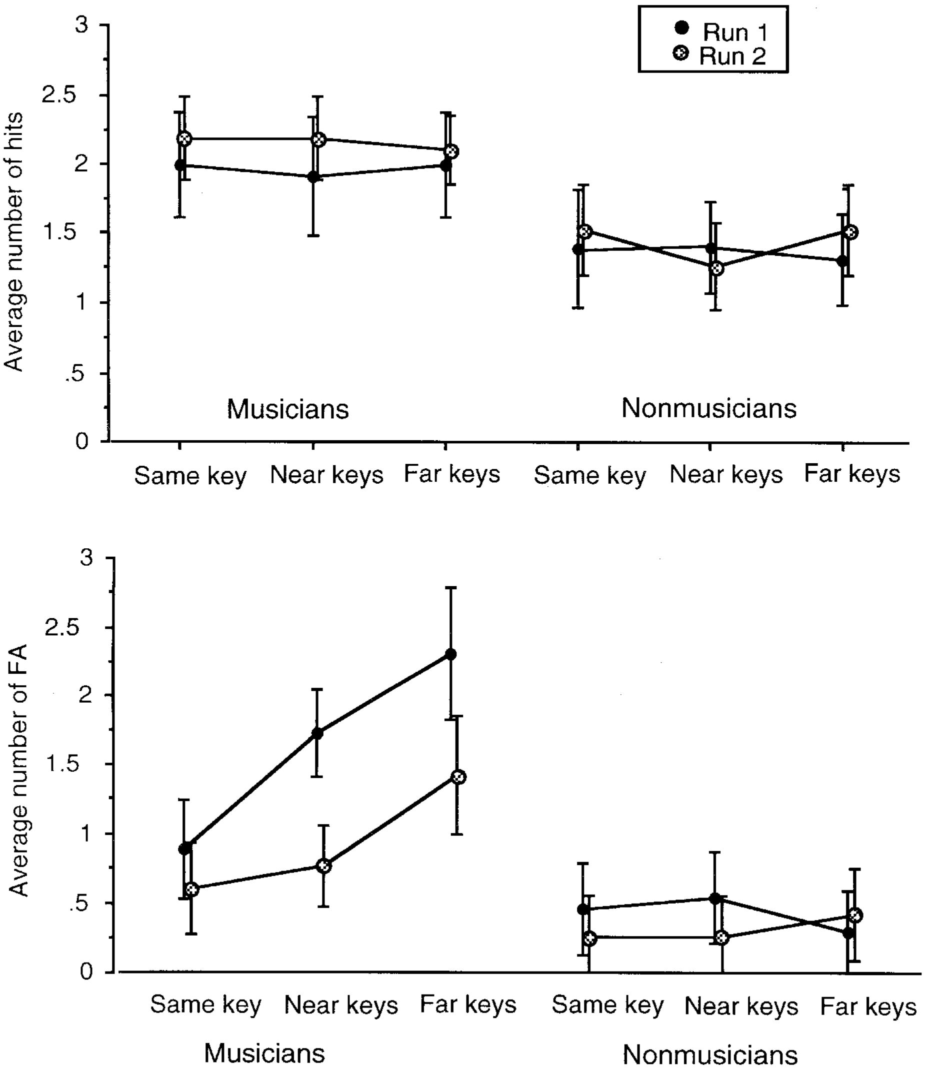

Figure 4. The 95% confidence intervals of average number of hits (top) and false alarms (FA) (bottom) as a function of key relatedness observed during the first and second runs in Experiment 1.

alarms were associated with the presence of intervals that were found by Kameoka and Kuriyagawa (1969) to be more dissonant when produced with complex tones (minor 2nd, augmented 4th, minor 6th, minor 7th, major 7th, minor 9th, and major 9th).

TABLE 1
Percentages of false alarms for different musical intervals[a]

Interval	*%FA*	*Interval*	*%FA*
Unison	0.00	Octave	2.39
Minor 2nd	4.79	**Minor 9th**	18.56
Major 2nd	0.00	**Major 9th**	8.98
Minor 3rd	1.80	Minor 10th	2.39
Major 3rd	1.20	Major 10th	0.60
Perfect 4th	3.60	perfect 11th	2.39
Augmented 4th	4.79	**Augmented 11th**	1.80
Perfect 5th	1.20		
Minor 6th	8.98		
Major 6th	0.00		
Minor 7th	19.76		
Major 7th	16.77		

[a] Bold face type indicates dissonant intervals according to Kameoka and Kuriyagawa (1969).

Discussion

Experiment 1 provided evidence that dividing attention between two melodies of a polyphony remains very difficult: Although the nursery tunes were short and very familiar, musicians detected only half of the errors contained in the polyphony, and their performance did not significantly increase with the number of runs. For the nonmusicians, the task was so difficult that they seem to have simplified it into one of selective attending. This finding thus does not support a true divided attention model. As already reported by Kahneman (1973), dividing attention between simultaneous messages is not easier with music than with other event structures such as spoken language.

The critical finding of Experiment 1 was that manipulating the key distance had no significant effect on melodic error detection but exerted a significant influence on the number of false alarms. On the one hand, this finding is consistent with the key-distance effect reported by Gregory (1990) and Sloboda and Edworthy (1981): In a divided attention task, manipulating the harmonic proximity of the underlying keys of the voices significantly affected the listeners' behaviour. On the other hand, it suggests that key distance did not facilitate the perceptual tracking of the melodies, but tapped into another perceptual mechanism that was brought into play when participants tried to follow two voices at the same time.

Experiment 1 makes it possible to go one step further and to shed some light on the nature of this perceptual mechanism. As previously described, false alarms were strongly associated with the presence of dissonant intervals. One sensory aspect of dissonance may be conceptualized as the quality of roughness that results from the interaction of nearby component frequencies of two sounds (Kameoka & Kuriyagawa, 1969; Plomp & Levelt, 1965; Plomp & Steenecken, 1968). Noninteger relations between their fundamental frequencies generally result in greater roughness. It has been proposed by Wright and Bregman (1987) that this roughness quality may be greatly lessened when two sounds belong to distinct auditory streams. This should have been the case in the present study since the melodies were played about an octave apart. Furthermore, this sensitivity to dissonant pitch intervals is surprising since participants were not supposed to react specifically to dissonant intervals. High false alarm rates thus suggest that the divided attention task may have encouraged the participants to integrate both voices into a single complex stream (Figure 1e). To further this interpretation, a second experiment was run in which participants were required to detect melodic errors in a single voice (selective attention task). If the sensitivity to dissonant pitch intervals was related to the dividing of attention, musicians should no longer react to the increasing number of dissonant pitch intervals in the near- and far-key conditions. To some extent, the nonmusicians' data in Experiment 1 provided some support for this assumption if we accept the hypothesis that they had actually been doing selective attending during the experiment.

EXPERIMENT 2

Participants

Twelve new subjects participated in this experiment. All were professional musicians having graduated from French national and regional music conservatories. All had received intensive training for over 10 years in music theory, ear training, and instrumental performance, and eight were practising composers. They were all currently involved in professional musical activity.

Material and procedure

The material was identical to that in Experiment 1. The experimental procedure was also identical except that this time the participants were required to track melodic errors occurring in the lower melody while ignoring those in the upper melody (selective attending task). Their task consisted of pressing a keyboard key as soon as a melodic error was detected in the lower tune. The errors were still present in the unattended melody. The two experimental pairs of melodies were tested in separate blocks. Each pair was presented in all three key-distance conditions in two successive runs as in Experiment 1. The tune used as the upper tune for half of the participants was used as the lower tune for the other half. In addition, the order of presentation of key-distance conditions within a run and of melody pairs across runs was counterbalanced across the participants. It was verified that all participants were able to name the four melodies prior to the experiment.

Results

Figure 5 presents the average number of hits and false alarms (out of two) as a function of key distance. Irrespective of the key distance, the error detection rate remained moderately high (72%). A 2 × 3 (Run × Key

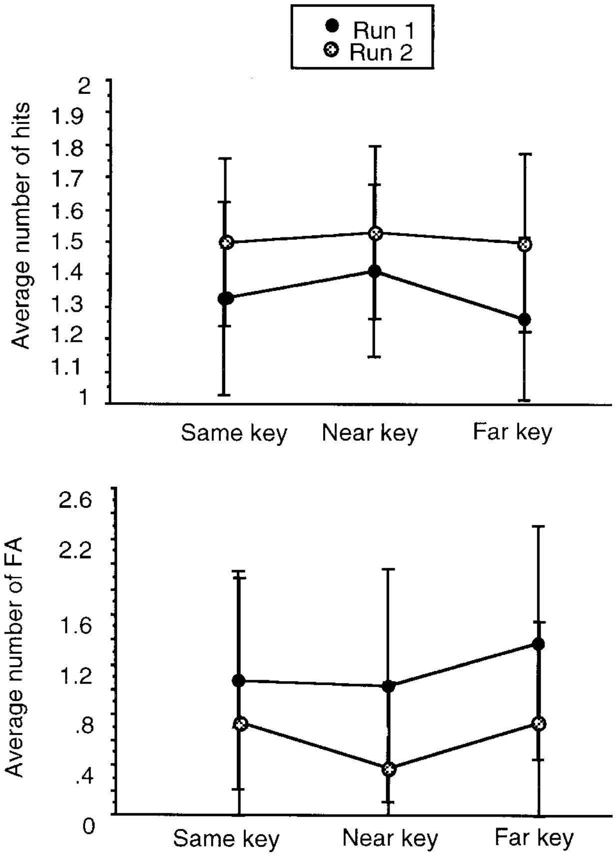

Figure 5. The 95% confidence intervals of average number of hits (top) and false alarms (FA) (bottom) as a function of key relatedness observed during the first and second runs in Experiment 2.

condition) ANOVA was performed with the two factors as within-subject variables. There was no effect of key distance nor of run on error detection rate. False alarms were less numerous in Experiment 2 than in Experiment 1, but this time their rate did not significantly increase with key distance.

GENERAL DISCUSSION

Previous research has provided evidence that listeners perform rather well in divided attention tasks with music (Gregory, 1990; Sloboda & Edworthy, 1981). The apparent ease with which several melodies at a time can be followed in music is consistent with the large audience for multivoiced music throughout the world and across different historical epochs. From a psychological perspective, it remained to be specified whether music makes possible a true dividing of attention or whether it enables a listener to simply develop perceptual strategies that compensate for the general limitations of attentional processes (as in Sloboda & Edworthy's figure-ground model). The present study was designed to test these two possibilities.

The results from Experiment 1 pose strong difficulties for a true divided attention model. In this experiment, the melodic errors introduced to highly familiar nursery tunes were all easily detected when the tunes were played one at a time. However, as soon as these tunes were played simultaneously, detection performance fell dramatically, even for musically expert participants. Playing these tunes in the same or in different (more or less related) keys had no effect on hit rates, further demonstrating the high difficulty of dividing attention in music[1].

On the contrary, the present finding provided several kinds of evidence that participants developed perceptual strategies to remedy the difficulty of dividing attention. Nonmusicians seemed to simply remove the difficulty by changing the divided attending task into one of selective attending. Musicians used a perceptual strategy that led them to produce numerous false alarms. The fact that these false alarms were mostly associated with dissonant intervals shed some light on this strategy. To be dissonant, the two pitches of a pitch interval need to be integrated into the same stream, at least partially (Wright & Bregman, 1987). The presence of false alarms suggests that participants focused on the concurrent organisation of the melodies and did not follow the two simultaneous messages individually by quickly switching attention back and forth between them (as illustrated in Figure 1d).

At first glance, the presence of numerous false alarms associated with dissonant intervals was rather surprising since the pitch range of the melodies should have been large enough (one octave on average) to remove any problem of stream segregation (although recent research using objective tasks such as melody recognition has shown an influence on performance of the presence of two streams of complex tones separated by as much as two octaves; Bey & McAdams, 1997). The fact that false alarms were mostly observed with a divided attention task (Experiment 1), but not with a selective attention task (Experiment 2), suggests that musicians tended to resolve the problem of dividing their attention by integrating the two streams into a single perceptual structure: Inharmonious pitch intervals present in a unified auditory stream would give rise to a perceptual attribute resulting from their interaction (i.e. roughness); this elementary auditory attribute is associated with musical dissonance, and unexpected dissonance in a tonal melody is often associated in turn with wrong notes, thus leading the musicians to press the key. This reasoning follows from much data that demonstrate that the computation of auditory attributes depends very strongly on the way the incoming sensory information is organized into events and streams. False alarms were more numerous in the far-key condition, because the key distance factor is necessarily confounded with the number of dissonant pitch intervals.

Integrating two (or more) voices of a polyphony into a single stream may be viewed as a perceptual strategy that can potentially compensate for the difficulty of dividing attention. A first obvious advantage is that a single, acoustically more complex, stream is easier to track attentionally than two more simple ones. Second, it is likely that once the residual harmonic quality of this single stream has been extracted, any change in one or the other melody would be easily noticed since it would modify the overall quality of the single stream. Sloboda and Edworthy's figure-ground model illustrates this second advantage: Once the listeners had learned the polyphony, they extracted its overall harmonic quality. Any change in the polyphony would be noticed by a simple comparison of the new quality of the single stream and the one stored in short-term memory. In this case, the integration of the polyphony into one stream would make it possible to attend to all the notes of both voices while focusing on only one "big" stream.

In contrast to Sloboda and Edworthy's study, we argue that this "integration strategy" is a general model for listening to multivoiced music. It does not only occur for Western tonal polyphony played in the same key: The fact that false alarms significantly increased with the key distance of the melody in the present study provides some evidence that even in these different key conditions, participants tended to integrate the voices into a single

---

[1] It may be argued that the null effect of key relatedness may be explained by the weak sensitivity of the experimental method. There are only a few observations per subject since only two pairs of stimuli were used, with four (Experiment 1) or even two melodic errors (Experiment 2) in each. We acknowledge that a greater number of observations per subject would strengthen the present conclusion. However, the null effect of key relatedness on the detection of melodic errors should be considered in the present study in light of the strong effect the same factor had on the false alarm rate. The specific effect of key relatedness on false alarms and not on hits suggests that key relatedness taps add a different perceptual processes than divided attention does.

perceptual structure. If the voices are integrated, their respective events interact perceptually in a way that does not occur when they are segregated. Moreover, it seems likely that this "integration strategy" is not limited (or specific) to Western tonal music, but may also occur for extra European multivoiced music or contemporary polyphonic music.

By contrast with the "figure-ground" model, an "*integration* model" would predict that the crucial factor affecting listeners' behaviour in a divided attention task relies on previous knowledge of the polyphony. When the listeners have never heard the polyphony before (as in the present study), they would have great difficulty distinguishing melodic errors from fortuitous dissonances in the polyphony (even if it is made of two very familiar tunes), because they cannot compare the new harmonic quality created by the melodic change to a previous well-known one. If, however, listeners had previously learned the polyphony, the detection of errors would be facilitated since they would notice the changes in the overall harmonic quality resulting from the introduction of melodic errors. It is likely that if the polyphony was learned in the same-key condition (as in Sloboda & Edworthy, 1981), error detection would be high. If the polyphony was learned in the far-key condition, error detection should be higher in the far-key condition than in the same-key condition. Finally, an integration model would predict that listening several times to a complete polyphony would facilitate learning by comparison with listening several times to each voice of the polyphony separately. Once again, this prediction should be observed whatever the keys of the individual voices.

Manuscript received April 1998
Revised manuscript accepted March 1999

REFERENCES

Bey, C., & McAdams, S. (1997). *Etudes des processus de formation des flux auditifs par une méthode objective.* Paper presented at the Actes du 4e Congrès Français d'Acoustique, Marseille.

Bigand, E. (1990). *Perception et compréhension des phrases musicales.* Unpublished doctoral dissertation, Université de Nanterre (Paris X).

Bregman, A.S. (1990). *Auditory scene analysis: The perceptual organization of sound.* Cambridge, MA: MIT Press.

Bregman, A.S. (1993). Auditory scene analysis: Hearing in complex environments. In S. McAdams & E. Bigand (Eds.), *Thinking in sound: The cognitive psychology of human audition* (pp. 10–36). Oxford: Oxford University Press.

Broadbent, D.E. (1958). *Perception and communication.* London: Pergamon Press.

Cherry, E.C. (1953). Some experiments on the recognition of speech, with one and two ears. *Journal of the Acoustical Society of America*, *25*, 975–979.

Dowling, W.J. (1973). The perception of interleaved melodies. *Cognitive Psychology*, *5*, 322–327.

Gregory, A.H. (1990). Listening to polyphonic music. *Psychology of Music*, *18*, 163–170.

Hirst, W., & Kalmar, D. (1987). Characterizing attentional resources. *Journal of Experimental Psychology: Human Perception and Performance*, *116*, 68–81.

Huron, D. (1989). Tonal consonance versus tonal fusion in polyphonic sonorities *Music Perception*, *9*, 135–154.

Huron, D., & Fantini, D. (1989). The avoidance of inner-voice in polyphonic music: Perceptual evidence and musical practice. *Music Perception*, *9*, 93–104.

Kahneman, D. (1973). *Attention and effort.* Englewood Cliffs, NJ: Prentice-Hall.

Kameoka, A., & Kuriyagawa, M. (1969). Consonance theory. Part I: Consonance of dyads. *Journal of the Acoustical Society of America*, *45*, 1451–1459.

Krumhansl, C.L. (1990). *Cognitive foundations of musical pitch.* Oxford: Oxford University Press.

Miller, G.A., & Heise, G.A. (1950). The trill threshold. *Journal of the Acoustical Society of America*, *22*, 637–638.

Neisser, U., & Becklen, R. (1975). Selective looking: Attending to visually significant events. *Cognitive Psychology*, *7*, 480–494.

Palmer, C., & Holleran, S. (1994). Harmonic, melodic, and frequency height influences in the perception of multivoiced music. *Perception and Psychophysics*, *56*, 301–312.

Peretz, I., & Morais, J. (1989). Music and modularity. *Contemporary Music Review*, *2*(1), 279–294.

Plomp, R., & Levelt, W.J.M. (1965). Tonal consonance and critical bandwidth. *Journal of the Acoustical Society of America*, *38*, 548–560.

Plomp, R., & Steenecken, H.J.M. (1968). Interference between two simple tones. *Journal of the Acoustical Society of America*, *43*, 883–884.

Sloboda, J., & Edworthy, J. (1981). Attending to two melodies at once: The effect of key relatedness *Psychology of Music*, *9*, 39–43.

Spelke, E., Hirst, W., & Neisser, U. (1976). Skills of divided attention. *Cognition*, *4*, 215–230.

Thompson, W.F. (1993). Modeling perceived relationships between melody, harmony and key. *Perception and Psychophysics*, *53*, 13–24.

van Noorden, L.P.A.S. (1975). *Temporal coherence in the perception of tone sequences.* Unpublished doctoral dissertation, Eindhoven University of Technology.

Wright, J.K., & Bregman, A.S. (1987). Auditory stream segregation and the control of dissonance in polyphonic music. *Contemporary Music Review*, *2*(1), 63–92.