Memory & Cognition
1997, 25 (4), 518-533

Recognition of music in long-term memory: Are melodic and temporal patterns equal partners?

SYLVIE HÉBERT and ISABELLE PERETZ
University of Montreal, Montreal, Quebec, Canada

The notion that the melody (i.e., pitch structure) of familiar music is more recognizable than its accompanying rhythm (i.e., temporal structure) was examined with the same set of nameable musical excerpts in three experiments. In Experiment 1, the excerpts were modified so as to keep either their original pitch variations, whereas durations were set to isochrony (melodic condition) or their original temporal pattern while played on a single constant pitch (rhythmic condition). The subjects, who were selected without regard to musical training, were found to name more tunes and to rate their feeling of knowing the musical excerpts far higher in the melodic condition than in the rhythmic condition. These results were replicated in Experiment 2, wherein the melodic and rhythmic patterns of the musical excerpts were interchanged to create chimeric mismatched tunes. The difference in saliency of the melodic pattern and the rhythmic pattern also emerged with a music–title-verification task in Experiment 3, hence discarding response selection as the main source of the discrepancy. The lesser effectiveness of rhythmic structure appears to be related to its lesser encoding distinctiveness relative to melodic structure. In general, rhythm was found to be a poor cue for the musical representations that are stored in long-term memory. Nevertheless, in all three experiments, the most effective cue for music identification involved the proper combination of pitches and durations. Therefore, the optimal code of access to long-term memory for music resides in a combination of rhythm and melody, of which the latter would be the most informative.

Recognition of familiar music is immediate and easy for everyone. Despite this apparent effortlessness, music recognition is a complex procedure that recruits multiple processing components. At the very least, recognition of a familiar piece of music entails a mapping procedure between the sensory input and a stored long-term representation that captures some of the invariant properties of the music. These long-term representations are assumed to be stored in a network that contains all the specific musical pieces to which one has been exposed during her/his lifetime. Given the size and variety of this network, it is essential to specify the nature of the sensory code that is used by the mapping procedure to select a particular stored representation.

There are a large number of sensory cues that can, in principle, serve as the entry code(s) to the network. Among these, the cues that arise from sequential variations along the pitch and the temporal dimension are apparently the most determinant. We will refer to these two classes of information with the generic terms *melody* and *rhythm*, as defined by pitch and temporal variations, respectively. Variations in intensity (dynamics), in spectral composition (timbre), and in speed (tempo) probably facilitate recognition insofar as these characteristics respect the structure of the original music or of its most frequently heard version. These variations are, however, considered secondary in that they are not determining factors of music recognition. For instance, most listeners easily recognize their national anthem, regardless of whether it is sung, played on an instrument lacking dynamics (such as the harpsichord), or played fast or slowly (within certain limits, as shown by Warren, Gardner, Brubaker, & Bashford, 1991). Thus, music recognition appears to rely most heavily on melodic and temporal cues. The question raised in the present study concerns their relative effectiveness as access codes to the musical stored representations.

This question grew out of neuropsychological observations. In a former study (Peretz et al., 1994), we found that brain damage can interfere with the ability to recognize music in an interesting way. Two patients (C.N. and G.L.), whose speech and cognitive functions were otherwise normal, were found to be unable to identify or experience a feeling of knowing for musical excerpts that were highly familiar to them before their brain injury. Prior knowledge of the music was checked with presentation of titles or lyrics that usually accompany the musical selections; this verbal information was readily identified by the 2 patients. Thus, both patients suffer from an auditory deficit that is specific to music. What intrigued us is that both patients were also found to be severely impaired in their perceptual analysis of sequential pitch variations. In contrast, none of them was impaired in discriminating and retaining

Order of authors is alphabetical since both authors equally contributed to the paper. This research was supported by a predoctoral fellowship to the first author and by a research grant to the second author, both from the National Science and Engineering Research Council of Canada. The authors are particularly grateful to Brian Ross, Marilyn Boltz, Andrea Halpern, and one anonymous reviewer for their constructive and insightful comments made on an earlier draft of this paper. Correspondence should be addressed to I. Peretz, Département de Psychologie, Université de Montréal, C.P. 6128, succ. Centre-ville, Montreal, Quebec, Canada H3C 3J7 (e-mail: peretzi@ere.umontreal.ca).

 Copyright 1997 Psychonomic Society, Inc.

temporal patterns. Thus, in principle, both patients could have used rhythm as a basis for recognition of familiar music. The fact that the patients could not recognize familiar tunes raised the possibility that melodic information, whose processing was impaired in both patients, is critical to access long-term memory for music.

The literature is scarce on this issue. To our knowledge, the only study that has dealt with the relative importance of pitch over temporal cues in the recognition of familiar music by normal listeners is that of White (1960). In that study, listeners were required to recognize familiar musical excerpts that had undergone several transformations. Among these, two are of relevance here. One transformation consisted of presenting the excerpts isochronously—that is, with every tone being of the same duration; only the melody was kept as in the original. The other transformation consisted of the reverse; that is, the excerpts were played on a single pitch, keeping only the original temporal pattern. When presented with the isochronous sequences, listeners achieved 88% correct identification. With the temporal pattern, they were 33% correct (chance level being 10%, as will be clarified below). This latter result clearly indicates that recognizing a well-known piece of music from its temporal structure alone is not straightforward, despite its wide acceptance as a reliable cue for music recognition (Dowling & Fujitani, 1971; Halpern, 1984; Lenneberg, 1967; Sturges & Martin, 1974). What this study clearly shows, however, is that melody is more determinant than rhythm in the recognition of familiar music. This outcome is thus highly consistent with the neuropsychological observations.

There are nevertheless a number of methodological points in White's (1960) study that deserve closer examination. One of these points is that top-down processing (i.e., from memory to perception), by way of consulting the stored representations to enhance recognition of the perceptual input, may have played a large role in White's results. Top-down processing was promoted by the fact that only 10 different musical selections were employed and that all 10 titles were provided as written multiple choices on the response sheet. Thus, subjects knew from the beginning of the experiment what the possibilities were. Moreover, each of these 10 excerpts was presented in 12 different formats along with its original version. Through repetition, even under substantial transformation, recognition may have been facilitated. Therefore, anticipation and prior exposure may have boosted performance to an artificially high level. Yet, it is unlikely that these factors fully account for the rather large discrepancy observed between recognition on the basis of the temporal structure and recognition on the basis of melodic structure; this difference appears genuine although it requires a more controlled assessment.

Reexamination of White's study was thus the primary goal of the present study. In particular, it is of interest to distinguish bottom-up processing, which is driven by sensory information, from top-down processing, which can arise from the activation of stored representations via conceptual information such as provided by titles. The idea is that, by distinguishing the data-driven and memory-driven sources of information, we will be able to better specify both the nature of the perceptual cues that are effective in activating the appropriate musical stored representation and the content of these memory representations. Given the paucity of research in the area, specifying the general conditions under which melody and rhythm contribute to the recognition of familiar music represents a much needed enterprise.

In effect, the primacy of pitch over time factors cannot be considered as a general property of music processing. This primacy can be reversed, depending on the experimental context. In judging similarity between pairs of brief unfamiliar musical sequences, subjects have been shown to exhibit higher reliance on temporal factors than on melodic ones (Monahan & Carterette, 1985). In judging whether or not a musical sequence was identical to a comparison sequence interleaved with distractor tones, subjects' attention was governed by rhythmic expectancies more than by melodic ones (Dowling, Lung, & Herrbold, 1987). These two studies illustrate that temporal information may take precedence over pitch information in some conditions. At the present stage, it is difficult to identify the parameters that influence the saliency of melody relative to rhythm in music processing. The experimental situations differ on too many levels, such as familiarity of the musical material, task demands, and musical background of the subjects, to draw general principles. By following up on White's study—examining more systematically the respective contribution of the melodic and the temporal structure to the recognition of familiar music—we aim to enlarge the database and, in doing so, provide a firmer ground against which the primacy of pitch over time factors can be weighted.

We explored this question in three experiments, using the same set of 25 familiar musical excerpts taken from popular songs. The selection of this particular subset was motivated by the need to collect naming judgments, which are among the best estimates of the precision of recognition. The stimuli correspond thus to the few musical excerpts that elicited name agreement in the vast majority of subjects in a former normative study (Peretz, Babaï, Lussier, Hébert, & Gagnon, 1995). Experiment 1 was basically a replication study of White (1960) with unanticipated and single presentation of stimuli. Subjects were presented with all 25 excerpts in either a melodic format, by being isochronous, or in a temporal format, by being produced on a constant pitch. Their task was to name, if possible, each excerpt and to rate their feeling of knowing it on a 5-point scale. In Experiment 2, the melodic pattern of one excerpt was combined with the temporal pattern of another excerpt so that each component was equally familiar but improperly matched. This procedure was used so as to achieve maximal comparability in isolating the temporal and melodic structure. Subjects were required to perform the same task as in Experiment 1, while focusing on melody and ignoring rhythm, or vice versa. Finally, in Experiment 3, a music-title verification task was employed instead of the naming task so as to measure the

possibility raised in the context of White's study that access to the title favors top-down processing.

EXPERIMENT 1

Experiment 1 compared melodic and temporal patterns as determinants of the first contact between an isolated and unanticipated familiar musical fragment and its representation in long-term memory. To do so, two groups of university subjects, selected without regard to musical training, were presented with either the melodic or the rhythmic version of 25 nameable musical excerpts. In the rhythmic group, subjects were presented with the temporal patterns devoid of pitch variations. In the melodic group, subjects were tested with the melodic patterns devoid of temporal variations. After the test, prior knowledge of the stimuli was verified with the original version of the tunes—that is, with their melodic and rhythmic patterns combined. For each stimulus, subjects were required to provide a naming response and a feeling-of-knowing judgment.

We take naming as the best indicator of the achievement of the mapping procedure between the stimulus and a stored memory representation, and we take feeling-of-knowing judgments as a less precise but convergent measure. Feeling of knowing was assessed by way of a 5-point rating scale, with 1 meaning *unknown* (i.e., no feeling of recognizing the stimulus) and 5 meaning *very familiar* (a high degree of confidence in recognizing it). These judgments usually refer to meta-memory phenomena (see Nelson, 1984, for a review). However, in the present study, these judgments are treated as an estimate of the retrievability of the target, following Koriat's (1993) conceptualization. These judgments are therefore considered as a comparative measure of the amount and intensity of activation of a stored representation that is elicited by hearing each musical excerpt when recovery of its title has failed. Numerous studies have shown that subjective conviction, as reflected by feeling-of-knowing judgments, is predictive of objective performance (e.g., Leonesio & Nelson, 1990). Therefore, in the present study, subjects who fail to retrieve the names of the presented music might nevertheless estimate with better-than-chance success whether they would be able to produce it in response to more information, as when presented with the original version of the music.

Method

Subjects. Thirty-two adults from the university community (range = 22–48 years of age; $M = 27.7$ years) volunteered to participate in this experiment. Since music is culture-dependent to an important degree, the subjects were required to have been living in Quebec since at least 2 years of age and to have French as their first language. They were recruited without regard to musical training. About a third of the subjects were categorized as musically experienced, having had 5 or more years of formal musical training (defined as lessons on an instrument or in music theory); none were professional musicians.

Stimuli and Material. The 25 familiar musical excerpts that served as stimuli (and are fully provided in Appendix A) were chosen from our pool of 144 musical excerpts normalized for familiarity and verbal associations (Peretz et al., 1995). The excerpts correspond to the musical part of folk songs and were rated as highly familiar, with a mean score of 4.9 and a standard deviation 0.36 on a 5-point scale, with 1 meaning *unknown* and 5 meaning *very familiar*. All stimuli were correctly named by at least 80% of the 60 subjects tested in the earlier study. No more than 25 musical selections fulfill this latter criterion. Since these judgments were obtained with computer-generated fragments, there was no need to verify the appropriate length, tempo, or instrumentation for full recognizability of the stimuli. Each fragment consisted of the basic melodic and temporal patterns as written in available scores. No dynamic, timbre, accompaniment, or performance timing variations were present. The fragments were one or two phrases long (according to Piston's, 1978, criteria), contained 10–30 notes ($M = 16.9$ notes), and lasted 3.9–14.5 sec ($M = 8.5$ sec). This version of the stimuli will be referred to as the *combined* version, since it incorporates both melodic and temporal variations.

Each excerpt was then submitted to two transformations that corresponded to the two different experimental versions. In one version, all pitch variations were removed and replaced by a single pitch that could vary from one excerpt to the next but that remained constant within the same excerpt (the pitches were selected from the middle range from C4 to C5). This transformation kept the original variations in duration and tempo; it will be referred to as the *rhythmic* version. In the *melodic* version, all temporal variations were removed, leaving only the original pitches. Each note duration was set to 600 msec, a value that is representative of a musical note duration according to Fraisse (1974). This procedure somewhat lengthened the stimuli in the melodic condition, with a mean duration of 10.2 sec. Musical notations for the three versions (combined, rhythmic, and melodic) as applied to "Bonne Fête" ("Happy Birthday") are displayed in Figure 1.

The stimuli were randomly sorted in each version and were generated by a Yamaha TX-81Z synthesizer controlled by an IBM-compatible computer. The analog output was preset to a MIDI-timbre imitating the sound of a pan flute and was recorded onto a cassette tape. Tapes were played back to the subject in free field at a comfortable listening level.

Procedure. The subjects were tested in small groups and were assigned randomly but in equal number to either the rhythmic or the melodic version. The subjects were first informed about the nature of the transformations applied to the sequences that they would be hearing. The session began with two examples, which had been taken from the same pool of familiar musical selections but had not been selected as stimuli. These two examples were transformed in the same way as the following experimental trials. The experimenter provided the correct response after each of these two examples and then told the subjects that they would no longer obtain feedback in the following experimental trials. The subjects were required to rate, on a 5-point scale, whether they had recognized the sequence (with 1 meaning *I did not recognize it at all*, and 5 meaning *I did recognize it very well*). They were further required to write down the title of the presented musical stimulus whenever possible. They had 6 sec to respond. After having been presented with the experimental version, the combined version was presented to them as a control for their prior knowledge of the stimuli. With the combined version, the subjects performed the same rating and naming tasks within the same time limits (i.e., 6 sec). The whole testing session lasted approximately 30 min.

Results and Comments

The subjects made few naming errors (corresponding to 12% of the total number of errors); most errors were omissions. These naming scores (with 1 for a correct title and 0 for an omission or incorrect title) as well as the ratings of knowing (from 1 to 5) were submitted to separate

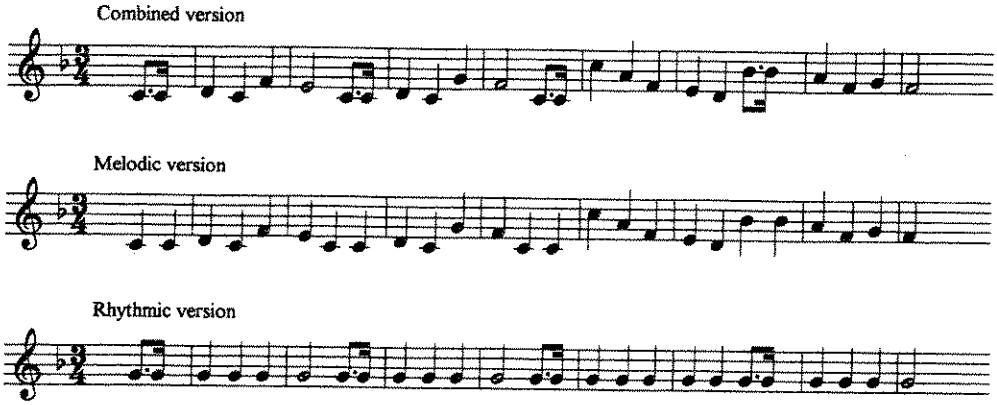



Figure 1. Example of stimuli used in Experiments 1 and 3. The combined, melodic, and rhythmic versions of "Bonne Fête" ("Happy Birthday") are displayed in the top, middle, and bottom musical scores, respectively.

analyses of variance (ANOVAs). Each type of response was analyzed by considering either subjects as the random variable ($F1$) or items as the random variable ($F2$). Group (rhythmic vs. melodic) and version (experimental vs. combined) were the between- and within-subjects factors, respectively, when subjects were the random variable. Group (melodic vs. rhythmic) and version (experimental vs. combined) were both treated as within-items factors when items were the random variable.

In Figure 2, the scores are represented in percentages of correct naming (left panel) and in mean ratings of knowing judgments (right panel) for each group being tested with either the rhythmic or the melodic version, as well as their results on the combined version. As can be seen, the two experimental groups did not differ in their naming scores of the combined musical excerpts, with means of 90% and 92% for the melodic group and the rhythmic group, respectively. Similarly, their confidence in knowing the combined excerpts was very high. As expected from our previous normative work, there was no noticeable difference between groups in prior knowledge of the musical material. All excerpts were easily named and well recognized. This was not the case for the melodic and rhythmic versions, as revealed by the interaction between group and version [$F1(1,30) = 80.73, MS_e = 6.34; F2(1,24) = 63.08, MS_e = 5.19$; both $ps < .001$]. Although the subjects generally performed more poorly with the experimental version than with the combined version ($p < .05$, by Tukey A post hoc comparisons), they produced more correct titles with the melodic version ($M = 49.2\%$) than with the rhythmic version ($M = 6\%$) [$t(30) = 7.348$, by an unpaired t test computed across subjects; $t(24) = 8.050$, by a paired t test across items; both $ps < .01$]. It is noteworthy that the correct naming responses obtained in the melodic version were distributed across 19 of the 25 stimuli. Six stimuli never elicited the correct title (for no obvious reasons), whereas the others did in a fairly distributed manner across both subjects and items.

The feeling-of-knowing judgments yielded a similar pattern of results (see right panel in Figure 2). The interaction between group and version was also significant [$F1(1,30) = 62.7, MS_e = 0.44; F2(1,24) = 68.93, MS_e = 0.34$; $ps < .001$]. As predicted, the ratings were reliably higher for the melodic version than for the rhythmic version [$t(30) = 9.746$, and $t(24) = 8.228$; both $ps < .001$]. Given the scale in use, with 1 meaning *unrecognized* and 5 meaning *well recognized*, a mean rating below 2.5 indicates low recognizability of the stimulus, whereas a score above 2.5 signals retrievability. Accordingly, rhythmic patterns of the familiar tunes were, on average, perceived as unfamiliar ($M = 1.6$), whereas the melodic patterns were judged overall to be familiar ($M = 3.5$).[1]

The performance obtained here with the combined version may have been enhanced by prior presentation in a different format, since all subjects were first tested with either the rhythmic or the melodic version of the same stimuli. To control for this possible bias, a new group of 15 subjects selected along the same criteria as in the present experiment were presented with the combined version only and were required to perform the same task. Essentially, the same high level of performance was obtained from these unprimed subjects. Their confidence in knowing the stimuli was extremely high ($M = 4.95$), and this was confirmed by their near-perfect naming scores ($M = 93.2\%$ correct). When the data obtained with this group are compared statistically with those of the subjects who performed the same tasks with the experimental versions, the same effects and interaction as those reported earlier were obtained. We can therefore confidently reject the possibility that the high levels of performance previously obtained with the combined version were due to repetition priming.

Musical training was not found to exert any significant influence in the present experiment. Number of years of instrument practice for each subject did not correlate with the ratings of feeling of knowing [$r(29) = .10$] or with the number of correct naming responses [$r(29) = .09$] across the two experimental versions.

In summary, the present experiment replicated the results initially reported by White (1960) with a larger sub-

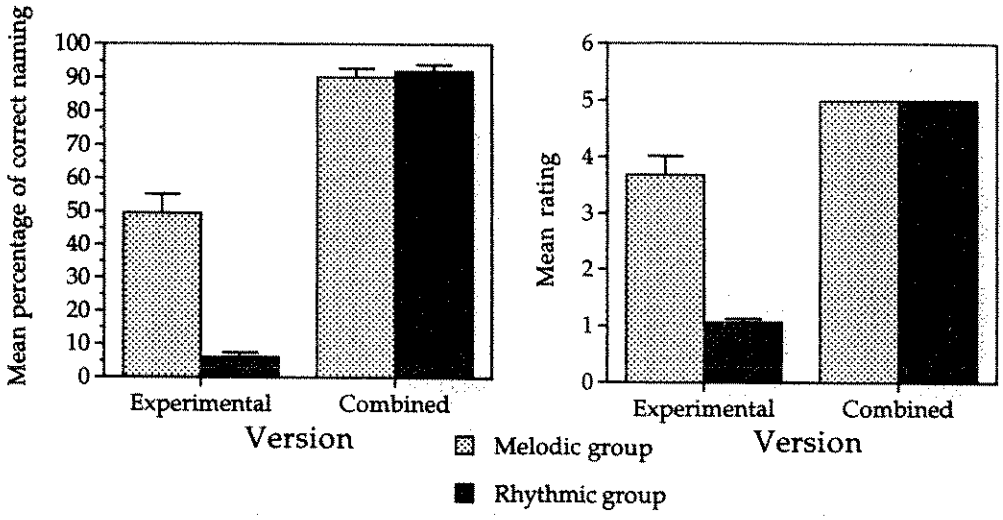


Figure 2. Mean percentages of correct naming (left panel) and mean ratings of knowing judgments (right panel) obtained by the melodic and rhythmic groups with the experimental versions and with the combined control condition in Experiment 1. Error bars represent the standard error of the mean (equal to 0.0 for knowing judgments in the combined version).

set of familiar and unanticipated tunes. The subjects here named about half of the musical excerpts on the basis of their melodic structure alone and only 6% on the basis of their temporal structure. Thus, the present results confirm the notion that familiar music is not easily identified by its rhythm. The melodic pattern is clearly more informative in this regard, although it does not fully account for music identification. Naming a tune on the basis of its melodic pattern does not reach the level obtained for the original excerpt where pitches and durations are combined.

EXPERIMENT 2

The results obtained in Experiment 1 may be conceived as arising from the way pitches and durations are differentially structured in popular songs. Popular music is very simple—particularly with respect to their temporal structure, which is basically organized around two durations (a short duration and a long duration in a ratio of 2 to 1). It is therefore possible that portions of the original rhythm of a tune incidentally coincide with an isochronous sequence. Such portions, which contain the original combination of melody and rhythm, may be sufficient for recognition. In other words, shifting tunes to a constant duration may be less detrimental to recognition than shifting them to a constant pitch, because the original tunes also contain a subset of notes with a constant duration but, less frequently, a series of notes with a constant pitch.

To reduce such coincidences, we presented here to new groups of subjects the same set of tunes with their respective melodic and rhythmic patterns interchanged (see Figure 3, for an example). With these mismatched tunes, each stimulus still possessed a melodic and a temporal pattern but came from different tunes. This procedure retains the same genre of musical structure and the same level of familiarity for the musical components while departing from the original arrangement of pitch and duration. Subjects were instructed to focus their attention on either the melodic or the rhythmic pattern of these chimeric tunes; otherwise, task demands were identical to those of Experiment 1. If recognition from melodic variations was indeed facilitated by isochrony in Experiment 1, then presenting the same melodic information with a new rhythm—that is, with temporal variations that are unrelated to the original tune—should considerably impair recognition performance.

Method

Subjects. Twenty-eight students (range = 19–42 years of age; $M = 25$ years) were selected using the same criteria as in Experiment 1. Six subjects had taken music lessons in the past (for 4–11 years); none were active performers.

Stimuli and Material. From the 25 musical excerpts used in Experiment 1, 19 could be paired in terms of note number and were kept as stimuli. With these 19 excerpts, we created 19 mismatched tunes by interchanging the melodic pattern and the temporal pattern. This was done by pairing two excerpts containing the same number of notes, without any concern for other musical parameters (see Figure 3). For 3 remaining tunes of the same length, the same procedure was applied so that none kept its original rhythmic pattern. All the mismatched tunes are provided in Appendix B.

Procedure. The same set of mismatched tunes served as stimuli in the two experimental conditions; only the instructions to the subjects differed. Half of the subjects were instructed to name and rate their feeling of knowing for the melodic pattern while ignoring the rhythmic variations. The other half were required to focus on rhythm while ignoring melodic variations. Otherwise, the procedure was identical to that of Experiment 1. The subjects were tested in small groups, with two examples followed by 19 experimental stimuli. After completing the task in the experimental condition, the sub-

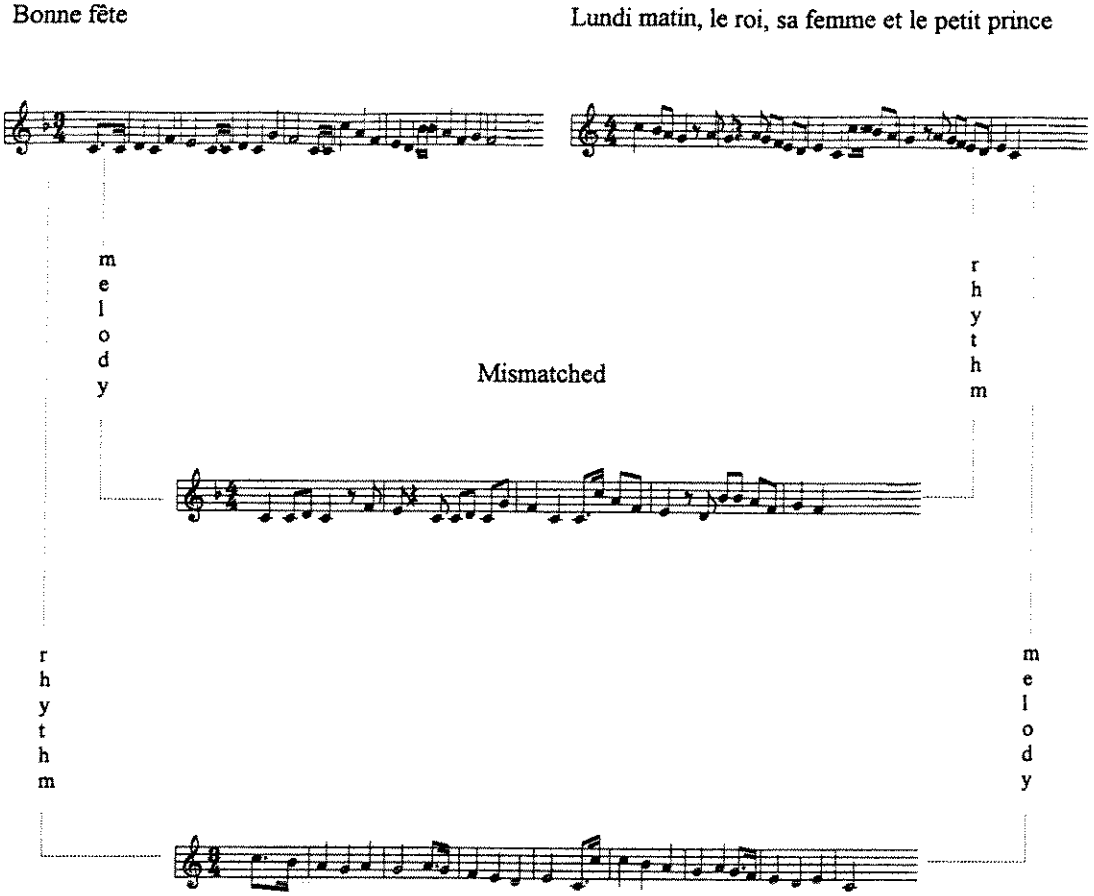



Figure 3. Example of stimuli used in Experiment 2. The original combined version of "Bonne Fête" and "Lundi Matin" are represented on top. Interchanging their respective melodic and rhythmic patterns produces the two mismatched tunes represented below.

jects were presented with the combined (i.e., matched) version of the musical excerpts.

Results and Comments

As can be seen in Figure 4, there is a striking similarity between the results obtained in this experiment and those of Experiment 1 (see Figure 2). Statistical analyses of the data were performed in the same fashion as in Experiment 1 and revealed a similar outcome.

First, the subjects were all highly and equally familiar with the musical material as indexed by their excellent performance with the combined (matched) version. However, this high rate of naming responses and of confidence level in recognizing the stimuli dropped with the mismatched tunes (i.e., the experimental version in Figure 4). On the naming scores with the mismatched tunes, the melody-instructed group was found to be clearly superior to the rhythm-instructed group ($p < .05$, by Tukey A post hoc comparisons). This pattern of results was supported by an interaction between group and version [$F1(1,26) = 31.78$, $MS_e = 7.30$; $F2(1,18) = 42.11$, $MS_e = 5.53$; both $ps < .001$]. As in Experiment 1, the subjects from the melody-instructed group named, on average, 53% of the stimuli on the basis of the melodic pattern. The rhythm-instructed group achieved only 8% correct (see left panel in Figure 4). Interestingly, the subjects who were instructed to focus on rhythm erred in giving the title corresponding to the melodic pattern. This occurred in 28% of their naming responses. It is as if they could not ignore the melodic information, which was even processed up to the level of name retrieval. These naming errors seldomly occurred in the melody-instructed group, with less than 2% of the trials. This rate of intrusions in the melodic condition was significantly lower than the one observed in the rhythmic condition [$t(26) = 4.119$, and $t(18) = 5.389$; $ps < .001$]. It is apparent that recognizability of the mismatched sequences was contaminated by the melodic dimension despite the fact that it had to be ignored. There is, thus, an asymmetry in selective attention, with attention being driven mainly by melodic information rather than temporal information.

The results on the feeling-of-knowing judgments mirror those of the correct naming responses (see right panel in Figure 4). The interaction between group and version was highly significant [$F1(1,26) = 39.60$, $MS_e = 0.18$; $F2(1,18) = 17.61$, $MS_e = 0.54$; both $ps < .001$], indicat-

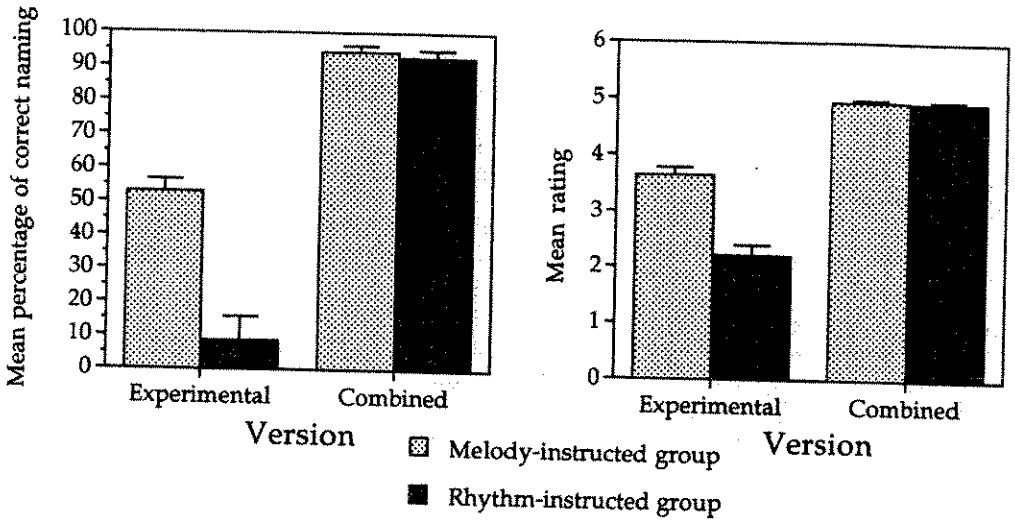


Figure 4. Mean percentages of correct naming (left panel) and mean ratings of knowing judgments (right panel) obtained for the experimental (mismatched) tunes by the melody-instruction group and the rhythm-instruction group, as well as their scores for the combined (matched) version in Experiment 2. Error bars represent the standard error of the mean.

ing that the subjects did not differ in their prior knowledge of the original tunes (i.e., in the combined version) but differed in response to the experimental version depending on the selective-attention instruction received. The experimental mismatched tunes were judged to be more familiar when the subjects focused on the melodic pattern than when they paid attention to the temporal pattern ($p < .05$, by Tukey A post hoc comparisons). In fact, the melodic pattern was generally judged to be familiar (with a mean rating of 3.7), whereas the temporal pattern of the same mismatched tunes was judged to be unfamiliar (with a mean rating of 2.2).

As in Experiment 1, musical training was found to have little impact on these tasks. In fact, the number of years of instrument practice was negatively correlated, but not significantly so, with both feeling-of-knowing judgments [$r(26) = -.29$] and correct naming responses [$r(26) = -.20$]. The absence of an expertise effect should, however, be interpreted with caution, since there were few musically trained subjects in the present samples.

It is still possible that recognition from the melodic variations occurred by coincidental matches with the unrelated rhythm. Indeed, it is plausible that the temporal structures of our stimuli are interchangeable because they are very similar. Hence, creating mismatched tunes would not alter the original musical structures to any significant degree. For the musically trained reader who can check the stimuli fully presented in the Appendices, this possibility is not very likely. The most fundamental characteristic of temporal structure—that is, metric organization—was changed from duple to triple, and vice versa, in more than half the stimuli. In such cases, the temporal accenting that occurs on a down beat falls at different points in time in the original and its mismatched counterpart.

We thought it desirable, however, to add some formal evidence in support of the notion that temporal structure was indeed modified by the mismatching procedure. To this end, we analyzed our two stimulus sets (the original excerpts and their mismatched formats) with the rules used by Boltz (1991, 1993), because she has addressed similar questions with stimuli coming from preexisting folk music. In her work, Boltz used basically two rules to define the sequence in terms of accent structure. One rule is related to melodic information in stating that tonic triad members (the tonic, mediant, and dominant) create melodic accents. The other rule is defined along the temporal dimension and attributes accents to notes that are prolonged in duration. When the two kinds of accents converge on the same note, then the joint accent structure draws attention to the same points in time and so facilitates processing and, hence, recognizability. We computed the serial position of these joint accents in our 19 original tunes and found that only 4 of the 18 that lent themselves to such an analysis kept these joint accents in the same serial positions in the mismatched tunes. Thus, for the large majority of our stimuli, we can conclude that changing the rhythm did modify its original accent structure. This modification of accent structure did not affect naming accuracy of the melodic component: The subjects identified more tunes when there was a disruption of the accent structure (with 59% correct) than when there was no obvious change (with 37%). Note, however, that the latter score concerns only 4 stimuli.

In summary, the present results clearly show that the melodic pattern is more indicative of identity than is its temporal pattern for highly familiar music. In the present experiment, where the melodic pattern was embedded in a new rhythm, there was no evidence of a decrement in

recognition and naming relative to when the melodic pattern was presented without rhythmic variations (in Experiment 1). The subjects were successful in ignoring the irrelevant rhythmic pattern, be it isochronous or unrelated to the original music. This was not the case for the temporal pattern, which remains a very poor cue for music identification.

EXPERIMENT 3

Compelling evidence for the more determinant role played by melodic information over temporal information in music recognition and identification has been gathered in Experiments 1 and 2. Melodic information does not, however, fully account for music recognition. In Experiments 1 and 2, the original combination of melody and rhythm was systematically better recognized and identified than was the melodic component considered separately. Thus, temporal information must contribute to the final decision, even if it is a very poor cue when considered in isolation. The level in the processing of the musical information at which temporal information contributes to the recognition process remains to be assessed. We hypothesized that perhaps rhythmic contribution to the recognition process lacks selectivity (to which we will refer as the *selection-difficulty hypothesis*). Temporal structures are shared, at least in large segments, across several tunes, as previously stated. Hence, temporal information may map onto the stored representation of multiple candidates and make response selection problematic. If correct, then narrowing the number of responses down to a single choice, by providing a title, should considerably enhance recognition performance on the basis of temporal cues. Temporal information could even reach the efficacy level of melodic information in a title verification task. Alternatively, temporal patterns may not contain sufficiently distinctive features to activate a single internal representation (to which we will refer as the *encoding-distinctiveness hypothesis*). In that case, providing a title for verification will not affect response decision; hence, recognizability will be low. Performance may, however, depend on top-down processing, as will be clarified below after the design of the experiment has been described.

The aim of Experiment 3 was to distinguish these two hypotheses with regard to the inferiority of rhythm relative to melody in recognition. The same set of tunes and transformations used in Experiment 1 served as stimuli here. There were two types of sequences: the melodic isochronous transformation and the temporal equitonal transformation. Again, the combined version served as a baseline against which to compare the respective contribution of melodic and temporal information. The major difference was the visual presentation of a title that corresponded to the auditory sequence in half of the trials. Titles were presented at different onset times with respect to the unfolding auditory stimuli. The title appeared at the onset, the middle, or the offset of the auditory sequence. Thus, there were nine experimental conditions, defined by the musical version presented (melodic, rhythmic, and combined) and the title onset asynchrony (TOA; onset, middle, and offset). Each condition was employed with a different group of subjects in order to avoid repetition priming.

TOA was manipulated here, enabling us to assess the contribution of top-down processing. Reading of a title is expected to activate one specific internal representation, which, in turn, will generate expectations that can be checked with the perceptual input, either directly by examining the auditory input or indirectly by affecting response decisions. If this top-down influence is direct, by supplying information to be checked in the unfolding sequence, then performance should be best when the title is provided early—that is, at the onset of the auditory input. When the title appears late, at the offset of the auditory sequence, the selection-difficulty hypothesis and the encoding-distinctiveness hypothesis entail opposite outcomes for the rhythmic patterns. Following the selection hypothesis, several candidates should have been activated at the offset of the auditory sequence, and the title would help to select one of them. In that case, performance should be fairly accurate. According to the encoding-distinctiveness hypothesis, the rhythmic sequence should not have activated a particular candidate in memory, and, hence, the late title appearance will not be effective as a selection cue. In that case, recognition performance should drop considerably. Presenting the title in the middle of the sequence was expected to produce intermediate effects. Because melodic pattern is believed to provide more distinctive cues for music identity, the title-verification task on the basis of melodic information should be less sensitive to the early versus late appearance of the title.

Since subjects were individually tested with on-line stimulus generation and data collection by computer, it was convenient to measure response times (RTs). The RT data were, however, expected to be less reliable than accuracy measures, given the small number of responses and the differences in time at which the target stimuli can be differentiated.

Method

Subjects. One hundred and thirty-five subjects (range = 18–51 years of age; M = 21.7 years), who did not participate in the preceding experiments, were tested here. They were selected using the same criteria as in Experiments 1 and 2. Eleven subjects considered themselves musicians (average number of instrument practice and of theory = 9.3 years; range = 3–16 years).

Material and Apparatus. The same three sets of 25 musical excerpts used in Experiment 1 (the combined version and their corresponding melodic and rhythmic transformations) served as stimuli. These stimuli were paired with their respective titles, containing 2–10 words (M = 4.2; see Appendix A), so as to create the experimental pairings. To create distractor pairings, 25 new pairs of tune and title were selected so as to be highly familiar (mean rating = 4.7 on a 5-point scale; range = 4.6–4.9; see Peretz et al., 1995) and to be similar in length to the experimental tunes. The distractor tunes lasted from 4.4 to 15.5 sec (M = 7.9 sec), and their titles contained 1–7 words (M = 3.7). The distractor tunes were transformed

in the same way as the experimental ones, so as to generate a combined version, a rhythmic transformation, and a melodic transformation. Unlike the experimental pairings, the distractor tunes and titles were randomly paired so as to avoid a true pairing. These true and untrue tune–title pairings were organized in three sets, each containing 25 experimental trials and 25 distractor trials. Thus, there was a combined set, a rhythmic set, and a melodic set corresponding to the combined, rhythmic, and melodic versions, respectively.

In order to reduce variability in RTs, the duration of each tune was kept constant across the three versions. To do so, the tone duration of the melodic version was determined by dividing the total duration of its combined version (as well as its rhythmic transformation, since they had identical durations) by the number of tones contained in the excerpt. Thus, unlike in Experiment 1 where tone duration was fixed at 600 msec, the tone duration of the melodic versions here varied across stimuli from 283 to 857 msec (M = 471 msec), while remaining constant within a melody.

Each set of stimuli was randomized and generated by the same equipment used in the previous experiments. The analog output was then digitized at a sampling rate of 44.1 kHz via the Soundtools program from Digidesign on a MacIntosh II FX computer and directly transmitted to the subject via a digital interface amplifier and Maico headphones. Coupling the visual presentation on the computer screen with the auditory stimulus, as well as with the response recordings, was done with the *Experimenter* program (Altmann, Wathanasin, Birkett, & Russell, 1992).

A trial consisted of the presentation of an auditory sequence synchronized with the presentation of a title on the computer screen. TOA was such that the title appeared (1) at the onset of the first tone of the auditory sequence, (2) in the middle of the sequence, corresponding to its total duration divided by 2 within 1-msec accuracy (on average, 4.25 sec after the beginning of the sequence), or (3) at the offset of the last tone of the auditory sequence (on average after 8.5 sec). These various TOAs are illustrated in Figure 5 with "Happy

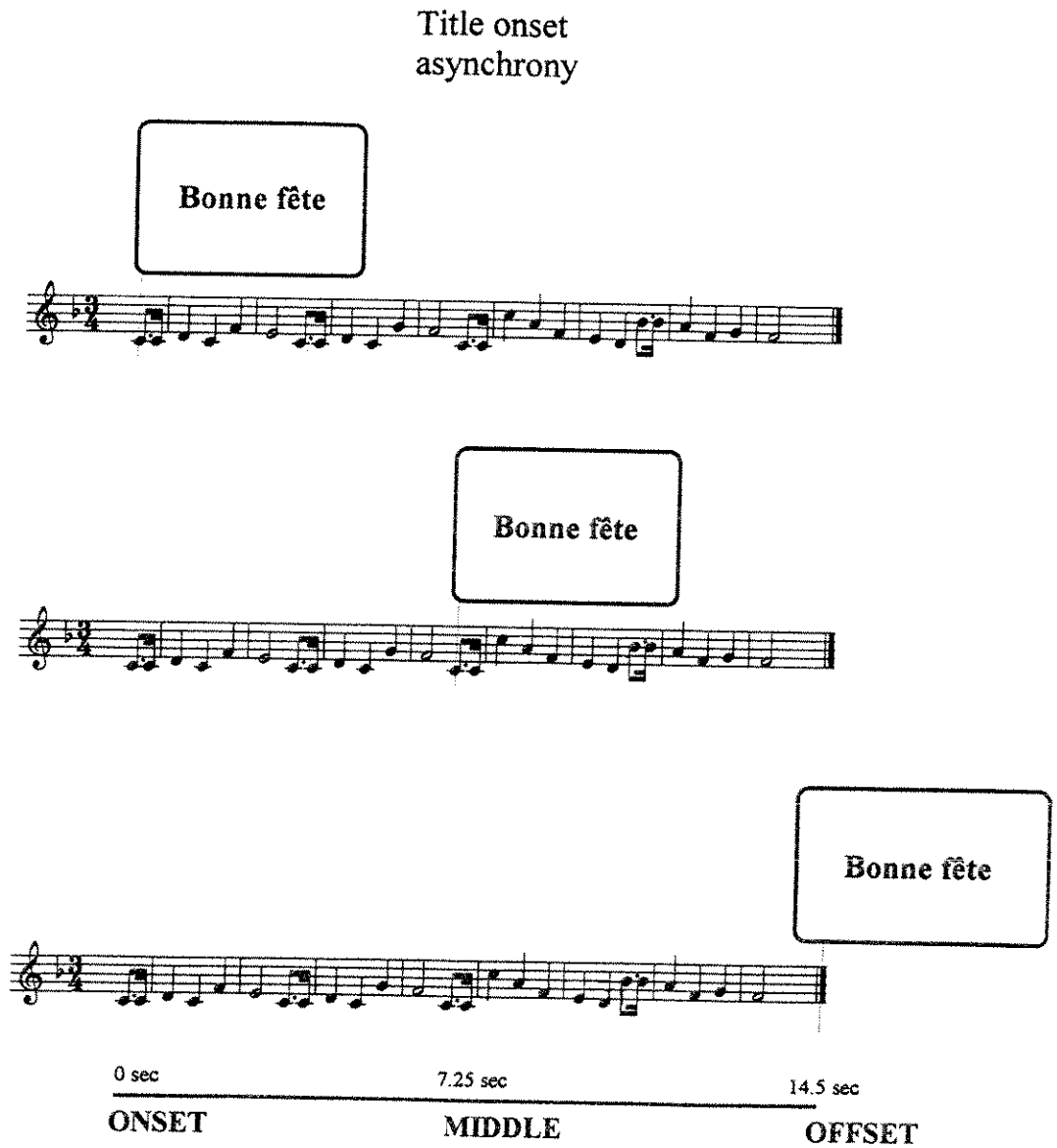


Figure 5. Illustration of the three different time asynchronies used in Experiment 3 between visual presentation of the title and the onset of the auditory presentation of "Bonne Fête" (in its combined version).

Birthday" presented in its combined version. Each set of stimuli was presented to the subject with a constant TOA. This created nine conditions, defined by the three musical versions (combined, melodic, and rhythmic) and the three TOAs (onset, middle, and offset).

Procedure. The subjects were randomly assigned but in equal number to one of the nine possible experimental conditions. There were thus 15 subjects performing the task under the same condition. The subjects were individually tested in a session lasting about 20 min. They were informed about the nature of the structural transformation applied to the musical sequences (if any) that they would be hearing. The experimental session began with four examples. The subjects were instructed to respond as quickly as possible whether or not the visually presented title corresponded to the musical excerpt that they were hearing. They responded "yes" or "no" on a response box that was connected to the computer. The RTs were computed from the onset of title presentation. The title appeared in the center of the screen and disappeared after 5 sec, or as soon as the subject had responded. The auditory sequence went on until its end and, 2 sec later, was followed by the next trial. No feedback was provided to the subjects.

Results and Comments

For each subject, the number of hits ("yes" responses for true pairings) and false alarms ("yes" responses for untrue pairings) were computed. The correction recommended by Snodgrass and Corwin (1988) was applied to these scores, by adding .5 to each value and dividing the sum by $N+1$, where N is the number of experimental and distractor trials for true and untrue pairings, respectively. The scores are presented in Table 1 along with the d' measure that takes both hits and false alarms into account. Statistical analyses were performed by way of a 3×3 ANOVA, taking d' as the dependent measure. The between-subjects factors were TOA (onset vs. middle vs. offset) and version (combined vs. melodic vs. rhythmic).

As can be seen in Table 1, title verification for the combined sequences was nearly perfect across TOAs. It was less accurate for the melodic version which remains, however, better recognized than the rhythmic version. In the latter case, a striking aspect of the results is that performance decreases systematically as title presentation is delayed. These trends were generally supported statistically.

Table 1
Mean Corrected Rates of Hits and False Alarms, Along With the Derived d' Scores, Obtained for Each Version at Each Delay (i.e., Title Onset Asynchrony) Between Auditory Presentation of the Sequence and Presentation of the Title

	Title Onset Asynchrony		
	Onset	Middle	Offset
	Combined Version		
Hits	.96	.96	.95
False alarms	.07	.06	.05
d'	3.43	3.46	3.45
	Melodic Version		
Hits	.90	.80	.80
False alarms	.28	.20	.15
d'	2.06	1.82	1.99
	Rhythmic Version		
Hits	.84	.75	.66
False alarms	.32	.40	.37
d'	1.60	1.00	0.83

There was a large effect of version [$F(2,126) = 178.50$, $MS_e = 0.34$, $p < .001$] and a main effect of TOA [$F(2,126) = 3.22$, $MS_e = 0.34$, $p < .05$]. However, the interaction between version and TOA just missed significance [$F(4,126) = 2.25$, $MS_e = 0.34$, $p < .07$].

Although the interaction between version and TOA failed to reach the conventional level of significance, it is apparent in the data that performance was differentially affected by the delay in title presentation in each version. To assess statistically the influence of title delay on performance, separate ANOVAs were performed for each version, considering TOA as the only between-subjects factor. These analyses revealed that increasing time delay between auditory presentation and title appearance decreased performance significantly in the rhythmic version only [$F(2,42) = 9.88$, $MS_e = 0.24$, $p < .001$]. In the combined and melodic versions, the effect of TOA was far from being significant (both Fs < 1).

Another aspect of the results that needed to be assessed separately was the apparent superiority of the melodic version over the rhythmic version at each TOA (see Table 1). These comparisons were assessed by unilateral t tests at each TOA. The melodic advantage was fairly general, although it was much reduced when title was presented at the onset of the auditory input [$t(28) = 2.00$, 3.54, and 6.27 at onset, middle, and offset TOAs, respectively; all ps < .05].

Thus, the accuracy measures show that title verification is generally an easy task to perform. Verification was better than chance in all conditions (chance level being 50% in accuracy and zero for d' measures), indicating that recognition can occur on the basis of rhythmic information alone. This conclusion could hardly be drawn from the results obtained in the two preceding experiments. In Experiments 1 and 2, rhythmic cues did not evoke a clear sense of recognizability. Another important result obtained with the title-verification task is that recognition accuracy was modulated by the moment at which the title appeared on the screen. Performance was best when title appeared early in the evaluation of the sensory input, particularly so with the rhythmic version. This effect of TOA on performance is highly consistent with the top-down hypothesis, whereby title presentation activates the corresponding musical stored representation, which, in turn, can supply the relevant information to be checked in the sensory input. When a title is presented at onset, the activated mental representation can be directly compared to the unfolding stimulus. When a title is presented at offset, such a direct verification of the sensory input can no longer take place; the title mainly serves to select the correct response. Recourse to direct verification of the sensory input seems to account largely for the results observed in the rhythmic version and less so for those in the melodic and combined versions.

The RTs provide support for the use of a verification procedure, with all musical versions. As can be seen in Figure 6, the subjects were slower to respond when the title appeared at the beginning of the sequence than when it appeared later. This indicates that the listener kept monitoring the input as long as it was available. In other words,

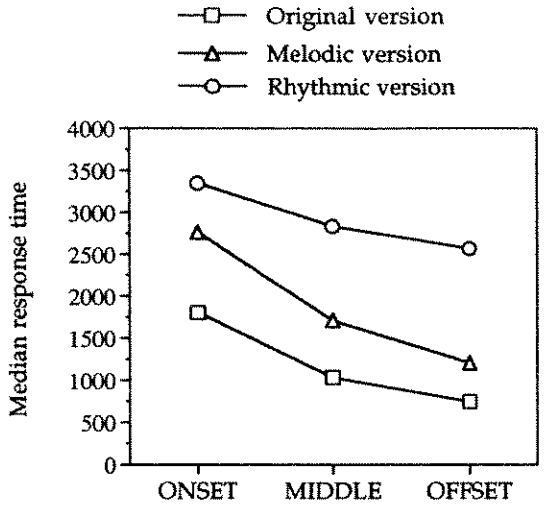


Figure 6. Median title-verification times as a function of the rhythmic, melodic, and combined auditory versions of the musical excerpts at the three TOAs in Experiment 3.

the more that information could be checked in the unfolding sequence, the more the subjects did so, and the longer were their RTs. This is a sensible strategy when the title appears at the onset of the sequence; the subject has to wait until he/she has gathered sufficient auditory information to respond. The use of this strategy is supported by a significant correlation between melody length and RTs [$r(24) = .63, .57$, and $.53$, for the combined, melodic, and rhythmic versions, respectively; $ps < .01$]. The fact that RTs still decrease from the middle to the offset of the sequence suggests that the subjects were careful to pick up as much information as they could.

Overall, the use of this serial strategy is consistent with the effect of TOAs observed in accuracy: The longer the RTs, or the more information there is to check, the more accurate the decision. This pattern should not, however, be taken as an indication of a speed–accuracy tradeoff. There was no positive correlation between speed and accuracy of the responses; rather, there were negative ones that support the convergence of the two measures: shorter RTs correlated with higher hit rates at each TOA [$r(43) = -.51, -.51$, and $-.40$ at onset, middle, and offset, respectively; all $ps < .005$]. Finally, the same hierarchy as that observed in accuracy was also obtained on RTs: The subjects performed fastest with the combined condition, a little slower with the melodic condition, and still slower with the rhythmic condition.

These main effects were each supported statistically. There was a robust TOA effect [$F(2,126) = 20.39$, $MS_e = 741{,}287$, $p < .001$], and there was a reliable effect of version [$F(2,126) = 45.53$, $MS_e = 741{,}287$, $p < .001$]. No interaction between these two factors emerged ($F < 1$). Planned comparisons revealed that the subjects were quicker in the melodic version than in the rhythmic one at middle and offset TOAs [$t(28) = 4.06$ and 2.87, at middle and offset TOAs, respectively; both $ps < .05$], but not at onset [$t(28) = 1.66$, $p > .05$].

There was one aspect of the RTs that puzzled us: Since naming the combined sequences was easy for every subject, as shown in Experiments 1 and 2, why did the subjects take so long to verify the title of the combined stimuli (i.e., 746 msec) when it appeared at the end of the sequence? At that point in time, most subjects should have already retrieved the title on the sole basis of the auditory input. It occurred to us that perhaps reading the title, per se, was time consuming. We tested 10 new subjects with the 25 experimental titles mixed with 25 new titles coming from unfamiliar folk tunes. The subjects had to decide whether or not the visually presented title was familiar. We found that the median RT to verify that the 25 target titles were indeed familiar was 900 msec. This is an estimate of the time required by the subject to read the target titles and to provide a familiarity judgment, which is a somewhat simpler decision than the one involved in the verification task used here. The 750-msec median RT observed here in the tune–title-verification task may reflect the time taken to read part of the title and to confirm one's internal decision.

In summary, the present results provide two new pieces of information with regard to the recognition process that takes place for familiar musical excerpts. First, access to long-term stored representations of the music via title presentation provides an important source of top-down processing that can be used to evaluate the sensory input. This top-down information is apparently used directly to check the auditory information. The fact that recognizability of rhythmic structure was particularly enhanced by this source of knowledge informs us that temporal structure is contained in the stored representations of the familiar musical excerpts. Another informative aspect of this result is that recognizability of temporal structure depends, to a large extent, on this top-down monitoring of the input. When the sensory input is no longer available for verification (i.e., when title appears at the offset), recognizability drops considerably. This latter result suggests that rhythm lacks distinctiveness at encoding rather than at retrieval. If, as we hypothesized, the lack of distinctiveness for rhythm were lying at the level of retrieval, due to the activation of too many compatible representations in memory, providing a single choice via title presentation should have been sufficient to narrow down the selection process even when title is presented at offset. The fact that performance dropped instead to a low level of performance (with 66% hits and 37% false alarms) confirms that rhythmic structure is a poor index for activating a stored representation by itself.

DISCUSSION

This series of experiments has provided compelling evidence for the predominance of melodic cues over rhythmic ones for providing the access code to the internal representations of musical pieces known since childhood. Melodic information appears to offer the minimal amount of information that can activate a stored representation in memory. Nevertheless, the optimal code seems

to require a combination of both melodic and rhythmic information.

The empirical evidence for the predominance of melodic cues over temporal cues was gathered in naming responses, feeling-of-knowing judgments, and speeded title verification. In Experiment 1, naming responses and feeling-of-knowing judgments revealed that transformations of familiar musical excerpts that kept the original pitch variations were clearly more diagnostic of music identity than were the original temporal variations. For about half of the musical excerpts, melodic information alone led to title retrieval. In contrast, the original temporal variations were judged to be unfamiliar and largely failed to evoke a title. This large advantage for the melodic pattern over the temporal pattern to activate music memories was reproduced in Experiment 2 with chimeric musical sequences. The latter were created by combining the melodic pattern of one familiar excerpt with the temporal pattern of another familiar excerpt. The subjects' task was to selectively attend to melody and ignore rhythm, or vice versa. They were found to weigh pitch structure more heavily than temporal structure, even when instructed to ignore the pitch variations. Generally, the subjects were not affected by what happened to rhythm when focusing on melody: They simply ignored temporal information. Finally, in the title-verification task used in Experiment 3, the subjects were found to match more quickly and, in general, more accurately the visually presented title with the auditory sequence when the latter contained melodic cues than when they carried only temporal cues.

We do not claim, however, that temporal structure plays a negligible role in music recognition. In each experiment, performance on the original tunes, which combine the melodic and temporal structures, was systematically superior to performance on the separate structures. Melodic information in isolation is often not sufficient to produce a perfect match with a stored representation; temporal structure must be included to do so. Furthermore, as shown in Experiment 3, temporal structure alone can lead to recognition when a title is provided as an early aid to monitor the sensory input. That is, the subjects were found to be quite accurate in matching a rhythmic sequence with a title while listening to the stimulus. However, when the title appeared too late, relative to the monitoring of the input (i.e., when the auditory stimulation had ceased), performance was found to be rather poor. The latter result provides important information with regard to the nature of the contribution of rhythm to the process of recognizing familiar music. First, rhythmic information must be contained in the stored representation of these musical pieces. Otherwise, there would be no possibility for the listener to use a top-down monitoring procedure. Second, the weaker contribution of rhythmic pattern to the recognition process is not simply due to a selection difficulty, following the idea that temporal patterns are shared to an important degree by many different tunes. If this were the case, title presentation by narrowing down the possibilities to a single choice should facilitate response decision, even when the title appears well after auditory presentation. The results are more compatible with the idea that, relative to melody, rhythm lacks encoding distinctiveness when the purpose of the task is to activate a representation in long-term memory.

The present results raise the question as to the degree of autonomy that melodic organization holds with respect to temporal organization in the recognition of familiar music. Are these two types of organization distinct, such that one organization is more effective than the other? Or should we consider that melody and rhythm are normally treated as a single entity, whose melodic content would be most salient? This question corresponds to a much debated issue in music cognition.

There are currently two distinct positions with regard to the question of autonomy for melody and rhythm. Tenants of the integration view, mainly developed by Mari Riess Jones and, more recently, by Marilyn Boltz, argue that melody and rhythm are treated as a unified dimension (e.g., Jones & Boltz, 1989). In their model, referred to as the *dynamic* model, temporal structure is considered to be the vehicle that drives attention toward or away from important melodic (i.e., nontemporal) information. This time-oriented strategy not only facilitates initial encoding but is believed to apply to remembering as well. Basically, it is suggested that pitch and time function as integral parts and, therefore, that remembering a melody involves the reinstatement of the original pitch–time structure. Tenants of the opposite view, the *independence* view, argue instead that listeners' percept of musical sequences results from an independent influence of melodic and temporal factors (Monahan & Carterette, 1985; Palmer & Krumhansl, 1987a, 1987b), at least in the initial stages of processing (Peretz & Kolinsky, 1993; Peretz et al., 1994; Thompson, 1994).

The studies that have addressed these two models empirically have yielded mixed results. In support of the independence view are findings that rhythmic patterns that conflict with melodic structure yield a much poorer performance than when the two coincide. Such integrality effects have been observed in a variety of contexts, including written recall (Boltz, 1991; Boltz & Jones, 1986; Deutsch, 1980), delayed recognition (Jones & Ralston, 1991; Jones, Summerell, & Marshburn, 1987; but see Smith & Cuddy, 1989, for negative results), retrospective duration judgments (Boltz, 1995), *same–different* classification judgments (Boltz, 1993; Jones, Boltz, & Kidd, 1982; Peretz & Kolinsky, 1993), and phrase-completion judgments (Boltz, 1989a, 1989b, 1989c; Schmuckler & Boltz, 1994). Moreover, interactive effects of rhythm and melody manipulations have been observed in selective-attention tasks; when subjects are explicitly instructed to ignore the rhythmic dimension (Bigand & Pineau, 1996; Boltz, 1989a; Kidd, Boltz, & Jones, 1984), they still exhibit response patterns that are affected by this dimension and conversely for the melodic dimension (Jones, 1987; Pitt & Monahan, 1987).

The other studies that have obtained results that are more compatible with the independence view have also involved a large variety of tasks. Melodic and rhythmic

structures have been shown to exert independent influences in similarity ratings for unfamiliar melodies (Monahan & Carterette, 1985; Monahan, Kendall, & Carterette, 1987), comparison of simultaneously or consecutively presented pairs of musical segments (Thompson & Sinclair, 1993), phrase-completion judgments (Palmer & Krumhansl, 1987a, 1987b), and change detection in individual tones and patterns (Thompson, 1994). Thus, task factors cannot easily account for the discrepancy between these studies and those supporting the integration model.

One of us (Peretz, 1990; Peretz & Kolinsky, 1993) has provided neuropsychological evidence that melody and rhythm are dissociable after brain damage. As mentioned in the introduction, processing of melodic information can be impaired while processing of temporal information remains intact, and vice versa. In a detailed case study (Peretz & Kolinsky, 1993), the dissociation between melody and rhythm was replicated across different sets of material and was maintained in conditions that promoted integration in the normal brain (by using a Stroop-like design in a *same–different* comparison task of brief musical sequences). These results argue against the view that melody and rhythm are treated as a unified dimension throughout processing. At the same time, they support the view that integration takes place after early separation of the two dimensions (see also Thompson, 1994, for a similar view).

The present results are also compatible with the view that integration occurs late in the processing of music—that is, after the initial stages of perceptual analyses that are believed to be largely preattentive and mandatory. The present observations suggest that, between access to the appropriate stored representation and recognition responses, an assembly process needs to take place whereby the melody information and temporal information are combined.

What remains unknown, however, is at which processing stages the combination takes place and to what extent this combination process requires integration of melodic and rhythmic information into a unique code. That is, integration into a unique representation may be required for the mapping process between the perceptual input and the stored representation of the tune. In that case, the melodic and rhythmic perceptual properties of this unique trace should be accessible separately. Alternatively, the two kinds of structure may not be integrated; rather, they may function as distinct but highly interconnected representations of the same tune. In that case, integration might concern the retrieval of the memory trace, which would combine these two distinct codes for adequate musical recognition. The present data are compatible with both possibilities. Further work would be required to distinguish between these two possible classes of functional organization. Nevertheless, the present results highlight the need to incorporate in both accounts a weighting factor in favor of melodic information as the most salient dimension in long-term memory for tunes.

The ensuing question is obviously the generalizability of the present findings. In this study, we have examined a rather limited set of tunes. This selection was strict in that only nameable tunes could be used. This raises the issue as to what extent the results obtained with these highly familiar tunes are generalizable to other tunes that are less familiar and to what extent they are generalizable to music in general. This is obviously a question to which we cannot provide any direct response. This should be the goal of future investigations. A change in paradigm would be necessary, however, since naming or title verification with most musical selections cannot be taken as a reliable index of music identification.

We predict, however, that the present findings would generalize at least to recognizability of most music of the Western culture. Indeed, the superiority of pitch factors over time factors may just reflect the regularities exhibited by the music to which we have been exposed since an early age. The music of our culture is simply more elaborate in pitch than in duration, and it uses pitch categories that are fixed to certain frequencies, whereas durational categories are more flexibly defined.

Western music is very stereotyped in rhythm. According to Fraisse (1982), who counted the frequency of different durations in various classical compositions ranging from Beethoven to Bartok, composers principally use two contrasting durations in any piece: a long duration and a short duration in a ratio of 2 to 1. This accounts for over 85% of the notes in a given composition. The large majority of the tunes (22 of 25) used in the present study fits with this description. Pitches are more variable. On average, the musical fragments contained 5.9 different pitches within the key of the tune. Yet, such variations in pitch can be considered as rather stereotyped because they do not include tonal deviations, as is generally the case in music compositions. Moreover, pitch was not harmonically elaborated here, as is also generally the case in most music. Nevertheless, even in these monophonic folk tunes, a disparity in variability along the pitch and duration dimensions is apparent and may account for their differential distinctiveness.

Furthermore, tones usually have a fixed pitch in our culture (with the A4 fixed at 440 Hz). Durations are less constrained, by allowing large deviations for the purpose of expressiveness (e.g., Gabrielsson, 1993). Perhaps, the rigidity of the pitch system affords efficient anchoring points for perception and memory, whereas the flexibility of the time structure serves mostly purposes other than recognition.

However, this asymmetry between melodic and rhythmic cues for recognition is likely to be culture-specific. Other cultures have more sophisticated rhythms (e.g., in ancestral Indian cults in the Martinique island, rhythmic patterns per se constitute musical tokens; see Desroches & Benoist, 1983). Even in less extreme instances, one would probably find cultures where a reverse contribution of melodic and rhythmic pattern to recognizability can be observed. With Balzano (1987), we might suppose that another culture whose music is less elaborate in pitch structure than ours would have a more rhythmically rich tradition as sources of additional structure in their music.

REFERENCES

ALTMANN, G., WATHANASIN, S., BIRKETT, T., & RUSSELL, P. (1992). *Experimenter* [Computer program]. Sussex, U.K.: Network Analysis Ltd., University of Sussex.

BALZANO, G. (1987). Measuring music. In A. Gabrielsson (Ed.), *Action and perception in rhythm and music* (pp. 177-199). Stockholm: Royal Swedish Academy of Music.

BIGAND, E., & PINEAU, M. (1996). Context effects on melody recognition: A dynamic interpretation. *Cahiers de Psychologie Cognitive*, **15**, 121-134.

BOLTZ, M. [G.] (1989a). Perceiving the end: Effects of tonal relationships on melodic completion. *Journal of Experimental Psychology: Human Perception & Performance*, **15**, 749-761.

BOLTZ, M. [G.] (1989b). Rhythm and "good endings": Effects of temporal structure on tonality judgments. *Perception & Psychophysics*, **46**, 9-17.

BOLTZ, M. [G.] (1989c). Time judgments of musical endings: Effects of expectancies on the "filled interval effect." *Perception & Psychophysics*, **46**, 409-418.

BOLTZ, M. [G.] (1991). Some structural determinants of melody recall. *Memory & Cognition*, **19**, 239-251.

BOLTZ, M. G. (1993). The generation of temporal and melodic expectancies during musical listening. *Perception & Psychophysics*, **53**, 585-600.

BOLTZ, M. G. (1995). Effects of event structure on retrospective duration judgments. *Perception & Psychophysics*, **57**, 1080-1096.

BOLTZ, M. [G.], & JONES, M. R. (1986). Does rule recursion make melodies easier to reproduce? If not, what does? *Cognitive Psychology*, **18**, 389-431.

DESROCHES, M., & BENOIST, J. (1983). Tambours de l'Inde à la Martinique: Structure sonore d'un espace sacré [Drums from India to Martinique Island: Sound structure of a sacred space]. *Études Créoles*, **5**, 39-58.

DEUTSCH, D. (1980). The processing of structured and unstructured tonal sequences. *Perception & Psychophysics*, **28**, 381-389.

DOWLING, W. J., & FUJITANI, D. S. (1971). Contour, interval, and pitch recognition in memory for melodies. *Journal of the Acoustical Society of America*, **49**, 524-531.

DOWLING, W. J., LUNG, K. M.-T., & HERRBOLD, S. (1987). Aiming attention in pitch and time in the perception of interleaved melodies. *Perception & Psychophysics*, **41**, 642-656.

FRAISSE, P. (1974). *Psychologie du rythme* [Psychology of rhythm]. Paris: Presses Universitaires de France.

FRAISSE, P. (1982). Rhythm and tempo. In D. Deutsch (Ed.), *The psychology of music* (pp. 149-180). New York: Academic Press.

GABRIELSSON, A. (1993). The complexities of rhythm. In T. Tighe & W. J. Dowling (Eds.), *Psychology of music. The understanding of melody and rhythm* (pp. 93-120). Hillsdale, NJ: Erlbaum.

GOODMAN, L. A., & KRUSKAL, W. H. (1954). Measures of association for cross classifications. *Journal of the American Statistical Association*, **49**, 732-764.

HALPERN, A. R. (1984). Organization in memory for familiar songs. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **10**, 496-512.

JONES, M. R. (1987). Dynamic pattern structure in music: Recent theory and research. *Perception & Psychophysics*, **41**, 621-634.

JONES, M. R., & BOLTZ, M. [G.] (1989). Dynamic attending and responses to time. *Psychological Review*, **96**, 459-491.

JONES, M. R., BOLTZ, M. [G.], & KIDD, G. (1982). Controlled attending as a function of melodic and temporal context. *Perception & Psychophysics*, **32**, 211-218.

JONES, M. R., & RALSTON, J. T. (1991). Some influences of accent structure on melody recognition. *Memory & Cognition*, **19**, 8-20.

JONES, M. R., SUMMERELL, L., & MARSHBURN, E. (1987). Recognizing melodies: A dynamic interpretation. *Quarterly Journal of Experimental Psychology*, **39A**, 89-121.

KIDD, G., BOLTZ, M. [G.], & JONES, M. R. (1984). Some effects of rhythmic context on melody recognition. *American Journal of Psychology*, **97**, 153-173.

KORIAT, A. (1993). How do we know that we know? The accessibility model of the feeling of knowing. *Psychological Review*, **100**, 609-639.

LENNEBERG, E. H. (1967). *Biological foundations of language*. New York: Wiley.

LEONESIO, R. J., & NELSON, T. O. (1990). Do different metamemory judgments tap the same underlying aspects of memory? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **16**, 464-470.

MONAHAN, C. B., & CARTERETTE, E. C. (1985). Pitch and duration as determinants of musical space. *Music Perception*, **3**, 1-32.

MONAHAN, C. B., KENDALL, R. A., & CARTERETTE, E. C. (1987). The effect of melodic and temporal contour on recognition memory for pitch change. *Perception & Psychophysics*, **41**, 576-600.

NELSON, T. O. (1984). A comparison of current measures of the accuracy of feeling-of-knowing predictions. *Psychological Bulletin*, **95**, 109-133.

PALMER, C., & KRUMHANSL, C. (1987a). Independent temporal and pitch structures in determination of musical phrases. *Journal of Experimental Psychology: Human Perception & Performance*, **13**, 116-126.

PALMER, C., & KRUMHANSL, C. L. (1987b). Pitch and temporal contributions to musical phrase perception: Effects of harmony, performance timing, and familiarity. *Perception & Psychophysics*, **41**, 505-518.

PERETZ, I. (1990). Processing of local and global musical information in unilateral brain-damaged patients. *Brain*, **113**, 1185-1205.

PERETZ, I., BABAÏ, M., LUSSIER, I., HÉBERT, S., & GAGNON, L. (1995). Corpus de 144 extraits musicaux: Indices de familiarité, d'évocation verbale et d'âge d'acquisition. *Canadian Journal of Experimental Psychology*, **49**, 211-239.

PERETZ, I., & KOLINSKY, R. (1993). Boundaries of separability between melody and rhythm in music discrimination: A neuropsychological perspective. *Quarterly Journal of Experimental Psychology*, **46A**, 301-325.

PERETZ, I., KOLINSKY, R., TRAMO, M., LABRECQUE, R., HUBLET, C., DEMEURISSE, G., & BELLEVILLE, S. (1994). Functional dissociations following bilateral lesions of auditory cortex. *Brain*, **117**, 1283-1302.

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

PITT, M. A., & MONAHAN, C. B. (1987). The perceived similarity of auditory polyrhythms. *Perception & Psychophysics*, **41**, 534-546.

SCHMUCKLER, M. A., & BOLTZ, M. G. (1994). Harmonic and rhythmic influences on musical expectancy. *Perception & Psychophysics*, **56**, 313-325.

SMITH, K., & CUDDY, L. L. (1989). Effects of metric and harmonic rhythm on the detection of pitch alterations in melodic sequences. *Journal of Experimental Psychology: Human Perception & Performance*, **15**, 457-471.

SNODGRASS, J., & CORWIN, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General*, **117**, 34-50.

STURGES, P. T., & MARTIN, J. G. (1974). Rhythmic structure in auditory temporal pattern perception and immediate memory. *Journal of Experimental Psychology*, **102**, 377-383.

THOMPSON, W. F. (1994). Sensitivity to combinations of musical parameters: Pitch with duration, and pitch pattern with durational pattern. *Perception & Psychophysics*, **56**, 363-374.

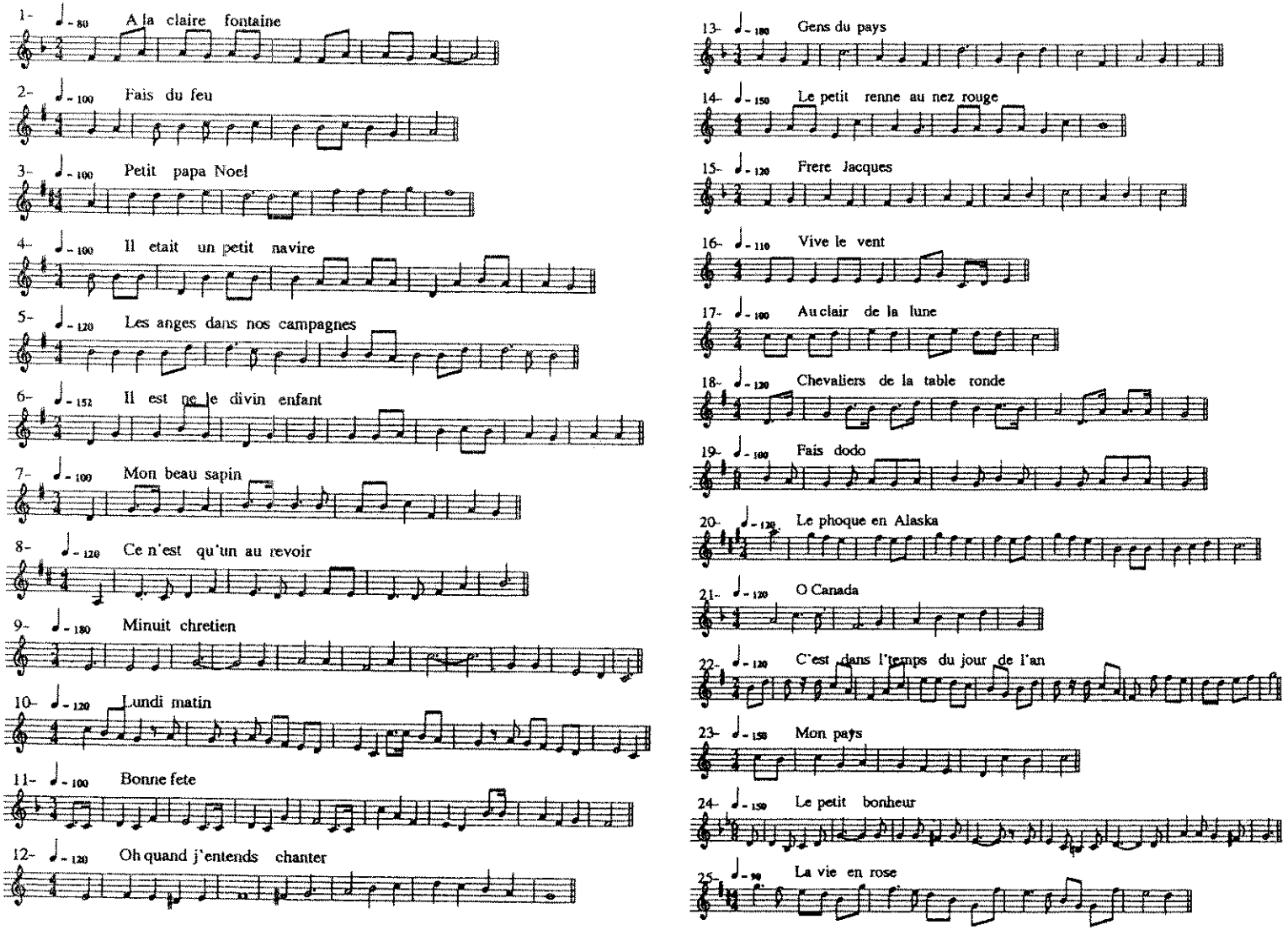
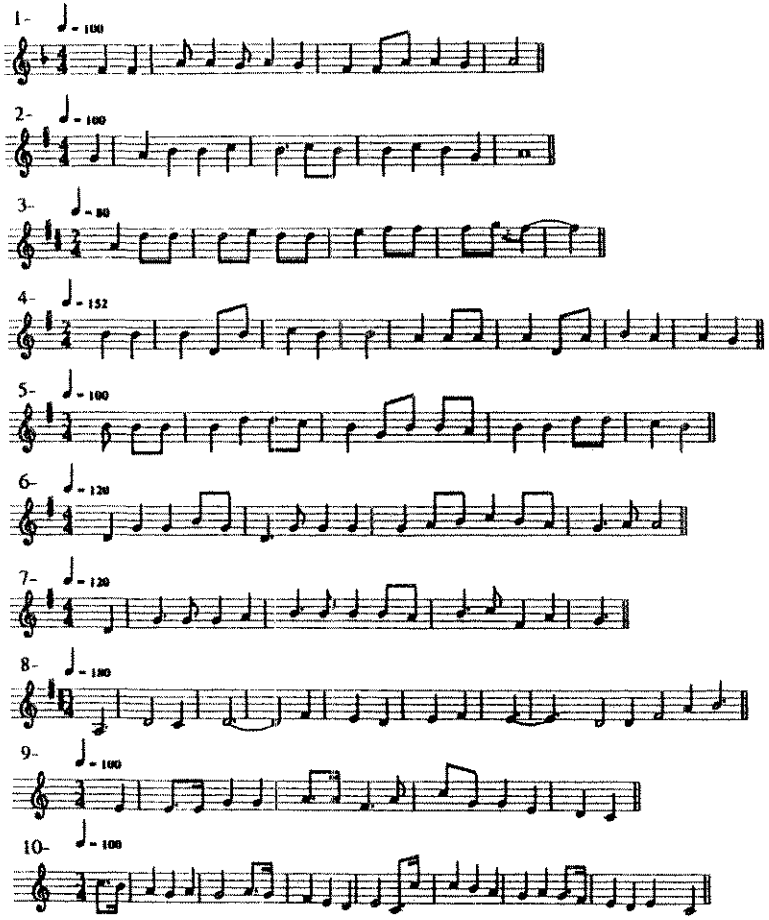
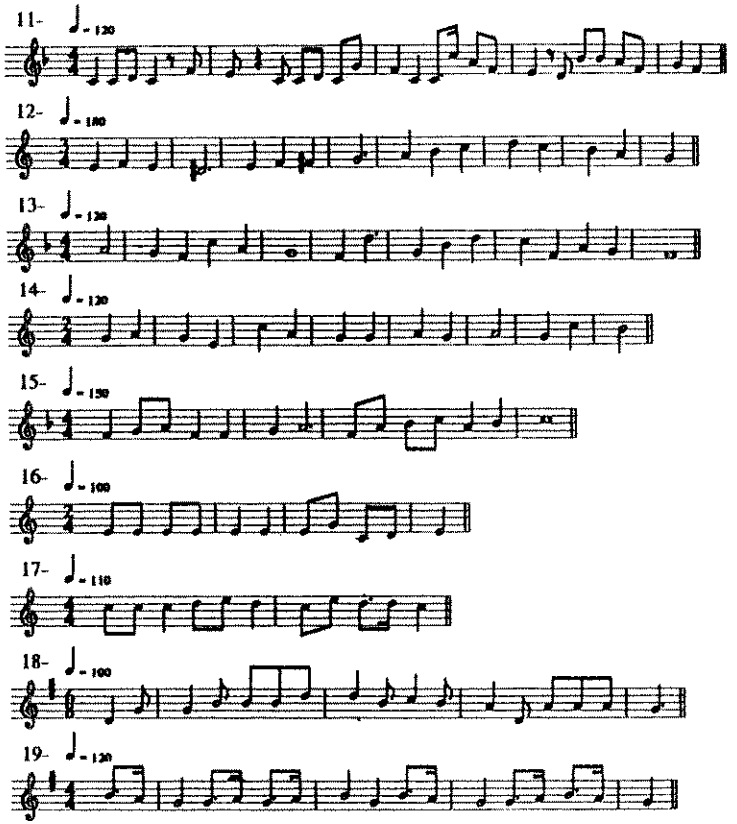
THOMPSON, W., & SINCLAIR, D. (1993). Pitch pattern, durational pattern, and timbre: A study of the perceptual integration of auditory qualities. *Psychomusicology*, **12**, 29-47.

WARREN, R. M., GARDNER, D. A., BRUBAKER, B. S., & BASHFORD, J. A., JR. (1991). Melodic and nonmelodic sequences of tones: Effects of duration on perception. *Music Perception*, **8**, 277-290.

WHITE, B. W. (1960). Recognition of distorted melodies. *American Journal of Psychology*, **73**, 100-107.

NOTE

1. The gamma correlations (Goodman & Kruskal, 1954) that are usually calculated between feeling-of-knowing judgments and subsequent recall of the solicited target information could not be computed because the subjects demonstrated near-perfect performance in the combined version.

APPENDIX A
Original Tunes Used in Experiments 1, 2, and 3

APPENDIX B
Mismatched Tunes Used in Experiment 2

(Manuscript received August 2, 1995;
revision accepted for publication August 6, 1996.)