ENCODING STRATEGIES FOR PITCH INFORMATION

Department of Psychology, Faculty of Letters,

Konan Women's University

1994

Mariko Mikumo

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGMENT

I wish to express sincere thanks and appreciation to Prof. Takao Umemoto for his timely and invaluable advice, both academic and personal, from the incubation period through the completion of this dissertation.

Many thanks are due to Profs. W. Jay Dowling, Diana Deutsch, Kengo Ohgushi, and Edward C. Carterette for their insightful comments and their careful, critical reading of the manuscript of my paper, when I submitted a part of this dissertation to *Music Perception*.

I would also like to express my deep gratitude to Profs. Jun–ichi Murai, Michie Doi, and Jun Kawaguchi for their support and encouragement while I was at Nara Women's University.

I also wish to thank Prof. Yoshie Yajima for her help in collecting musically highly trained subjects at Kyoto Women's University. I am also grateful to all subjects : students of Nara Women's University, Kyoto Women's University, Nara University of Education, Konan Women's University, Yamate Women's College, Kyoto University, Aichi Prefectural University of Fine Arts and Music, Doshisha Women's College, elementary schools, junior and senior high schools in Kyoto city, and members of Kyoto and Ashiya citizen orchestras, for their help.

The research of this dissertation was supported by the Research Fellowships of the Japan Society for the Promotion of Science, and by the Grant–in Aid for Scientific Research from the Ministry of Education, Science and Culture.

ABSTRACT

ENCODING STRATEGIES FOR PITCH INFORMATION

Mariko Mikumo

Department of Psychology, Faculty of Letters,

Konan Women's University, 1994

The purpose of this study was to investigate the encoding strategies for pitch information of short melodies. "Encoding" refers to the establishment of codes which are more likely to be stored in relatively permanent long-term storage. The encoding processes are often assigned in the short-term memory.

A variety of mental representation models have been proposed for pitch structures (see Hubbard & Stoeckig, 1992). Psychoacoustical models hypothesized that pitch could be represented by a single dimension (e.g., Stevens, Volkmann, & Newman, 1937). Rule-based models exploit some of the outward similarities between music and language (e.g., Deutsch, 1980 ; Deutsch & Feroe, 1981 ; Lerdahl & Jackendoff, 1983). In schematic models, pitch is represented by some sort of music schema which consists of at least three, or more dimensions (e.g., Krumhansl & Kessler, 1982 ; Lerdahl, 1988 ; Longuet-Higgins,

1978 ; Shepard, 1982a,b). In connectionist model, harmonic relationship (musical chords) are represented by nodes in a network (Bharucha, 1987 ; Bharucha & Stoeckig, 1986, 1987).

There are a lot of models of the mental representation for pitch structures, indicating that any given music stimulus is likely to have multiple representations of pitch, and that, therefore, people would employ multiple coding strategy in auditory modality for pitch sequences (West, Howell, & Cross, 1991). Various representations of pitch in these models are sophisticated theoretical and conceptual representation for pitch structure, but are not likely to be auditory imagery perceived from pitch information. Furthermore, it is possible that auditory pitch information is encoded not only in auditory modality, but also in visual or kinesthetic modality ; namely that multimodal representations based on intermodal coordination are employed to encode auditory pitch information. Therefore, in the present study, the encoding strategies for pitch information in short melodies were investigated, in terms of not rule- or schema-based representations but imagery-based multimodal representations.

Posner (1973) proposed concerning codes in memory that (a) there are at least three types of codes, visual, verbal, and motor (Bower, 1972b) ; (b) each code endures, and is not a transient residual of stimulation ; (c) people differ in their propensity to use each type of code ; (d) these codes are parts of separate memory systems that can be examined in isolation in the laboratory. In this study all four points were experimentally demonstrated.

"When you memorize and retain a melody, what types of strategies do you use to encode pitch information ?" The preliminary questionnaire data indicated that the subjects used one or more of the following several strategies : (a) a verbal encoding strategy, in which each pitch in a melody was labeled with the name of a musical note, and this code was rehearsed and stored in memory ; (b) a sensory (auditory) encoding strategy, in which pitches in a melody were retained in memory as auditory information ; that is, by singing, whistling, humming, mental rehearsal of pitches, and so on ; (c) a visualizing strategy, in which pitches were visualized in their image, as a melodic contour, on a keyboard, or on a staff notation ; or (d) a motor encoding strategy, in which an auditory melody was encoded by the movement of the fingers as if playing the piano. Some subjects reported that they used two or three strategies simultaneously. The questionnaire data presented several important implications, which were investigated in Experiments 1 to 9 in this study.

The findings of these experiment are described below.

When the subjects were asked to memorize and retain the pitch information in short melodies, the performance of the highly musically trained subjects was consistently better than that of the untrained subjects, because the former subjects could listen analytically to the musical series, with considerable attention to the internal relationships among their components (i.e., pitch interval), and because they employed their own strategy effectively or were motivated to try to employ several strategies at the same time to encode pitch information. Under ordinary conditions, their dominant and effective strategy of encoding the pitches of tonal melodies was verbal rehearsal of note names. When this verbal encoding strategy was employed, pitch or pitch intervals could be retained accurately for a long time (Experiment 1).

The neuropsycological evidence that the highly musically trained subjects use verbal encoding strategy for tonal melodies was found in Experiment 3–3, in which the tonal melodies, which are actually non–verbal stimuli in themselves but are "note name–evoking" stimuli, would be processed as verbal stimuli in the left hemisphere. The preferential left hemisphere lateralization found in the highly musically trained subjects would be due to the cognitive "linguistic" structure of tonal melodies, which implies the likelihood that pitches are encoded as verbal labels (note names) and that the processing of musical tasks involves sequential programs most analogous to those of language and speech. On the other hand, the atonal melodies would be processed as non–verbal stimuli in the right hemisphere. In Experiment 3–4, the stimuli were the melodies sung with note names at accurate pitches, so that the subjects were given the verbal codes at the same time and it was not necessary for them to encode pitches as verbal code (note names) by themselves. In this experiment, the highly musically trained subjects could retain the pitches not only in tonal melodies but also in atonal melodies as verbal codes ; when there was a possibility that the pitches were encoded as note names, these melodies were processed in the left hemisphere. However, the less well musically trained subjects could not encode pitches as note names, even when the pitches in a melody and their verbal codes were given at the same time, and they therefore processed tonal and atonal melodies as non–verbal stimuli in the right hemisphere.

The ability to encode pitch information as verbal labels would be closely related to the ability to detect a 50–cent deviated tone out of a melody, because the 50–cent deviated tone from an equal–temperament scale is the most difficult pitch to categorize into a note name on a chromatic scale. Although the ability depends considerably on experience or training in music, it was found that even those who were less well trained in music had acquired the ability to some extent as their age increased. Moreover, the later the deviated tone appeared in the melody, the greater the accuracy with which it was detected. In the detection of a deviated tone out of a melody, the subject gradually constructs an internal cognitive framework (scale schema) upon hearing the tones from the beginning of the melody. This cognitive framework plays an essential role in the perceptual interpretation of each succeeding tone, and the subjects can detect a deviated pitch out of a melody by referring to the framework. In this experiment, the subjects would constructed the diatonic scale schema upon hearing the tonal melodies, so that the detection of a deviated tone out of the tonal melodies was more accurate than out of the atonal melodies (Experiment 2).

Since it was difficult for the less well musically trained subjects to encode pitches as note names, they attempted to encode pitches as acoustic pitch codes through humming or mental rehearsal of pitches, and this was also true for the highly trained subjects with atonal melodies. In this case, they tended to listen globally to the pitch sequences on the basis of the total configuration ; the melodic contour was the dominant and effective cue for encoding pitch sequence (Experiment 1). Melodic contour is the other property as important as pitch or pitch interval in terms of melody recognition, and there is an assumption that various visual and auditory contours are perceived in broadly similar ways.

In Experiment 4, the auro–visual intermodal relationships based on visual representations were investigated. It was found that the less well musically trained subjects employed a visualizing strategy in which an auditory contour (auditory imagery) was visualized as a visual contour (visual imagery), and that visualization of auditory imagery as melodic contours would be to some extent an effective strategy for them to retain pitch sequences. Their internal visual representations of melodic contour roughly reflect the auditory pitch intervals. In stead of melodic contour, the highly musically trained subjects visualized pitch information as staff notations, in which notes appeared at accurate positions on a staff, especially for tonal melodies. That is, the visual distance between notes in their internal representations precisely reflect the auditory pitch intervals. Visualization of auditory imagery as staff notations would be an effective strategy to encode the pitches of melodies.

The evidence that the highly musically trained subjects precisely tracked their internal visual representations of notes on a staff, especially for tonal melodies, and that the less well trained subjects roughly tracked their internal visual representations of melodic contour, while employing visualizing strategy was obtained in the study of eye movements (Experiment 5).

In Experiment 6, it was found that subjects employ two or three codes at the same time rather than just one to memorize or retain pitches. For example, pitch rehearsal of auditory information along with note names (dual coding) and, to a greater extent, at the same time visualizing the staff notation (triple coding) were the most effective strategies for the highly musically trained subjects with tonal melodies. Pitch rehearsal along with visualizing the melodic contour (dual coding) was effective strategy for the highly musically trained subjects with atonal melodies and for the less well trained subjects with both tonal and atonal melodies.

In Experiment 7, the intermodal auro–motor coordination based on spatio–motor representations was investigated. It was found that piano majors effectively employed an external motor encoding strategy, in which auditory information was encoded as finger movements analogous to playing the piano. The motor encoding strategy was stable or robust against interference and time decay ; that is, the effects of finger movements became stronger as melody length and retention interval increased, and this was especially true when interference stimuli were interpolated during the retention interval, because the more challenging the situation became, the greater the efforts of the subjects to reduce the latency and to make elaborative and rapid finger movements. Some subjects could employ this external spatio–motor representation (motor encoding strategy) effectively on their own, while others could not employ it independently of verbal rehearsal of note names.

In Experiment 8, it was found that, when pitches of melodies were encoded, the visuo-spatial representation (i.e., visual image of spatial configurations) for pitch information were more accurate in the compatible spatial directions related to the motor system used in playing the instrument (internal spatio-motor representations). Piano majors would have spatial images in which the right and upper directions are compatible with higher pitch. For violin majors, the forward direction is compatible with higher pitch. For violoncello majors, the lower and backward directions are compatible with higher pitch. For vocal music majors, the right, lower and backward directions are compatible with higher pitch. For less well musically trained subjects, the upper direction is compatible with higher pitch.

The evidence that while encoding pitches subjects who are highly trained in playing the instrument precisely tracked their internal visuo-spatial representations in the compatible spatial directions related to the motor system used in playing the instrument was obtained in the study of eye movements (Experiment 9).

An attempt is made to interpret these results in relation to Paivio's dual-coding theory, the levels-of-processing theory, and Baddeley's working memory model.

As described in Chapter 1 and Experiment 6, Paivio (1969, 1971, 1978, 1986) proposed the dual-coding theory, the essence of which is that there are two basic independent but interconnected systems for the representation and processing of information. The verbal system deals with linguistic information and stores it in an appropriate verbal form, while the nonverbal system carries out image-based processing and representation. Within the two systems there are associative structures. Both systems have basic representational units that are linked to one another by referential connections : Logogens for the verbal system and Imagens for the nonverbal system. Both the Logogens and Imagens are further divided into sub–systems (i.e., visual, auditory, kinesthetic, gustatory, and olfactory) which process either verbal or nonverbal information in the different modalities (modality–specific).

Considering the findings in Experiments 1, 4, and 6, for example, the dual–coding by pitch rehearsal along with note names is consistent with Paivio's theory, because pitch rehearsal would be an auditory Imagen and note names would be a auditory Logogen. If the triple–coding by visualizing staff notation along with pitch rehearsal with note names is consistent with the Paivio's theory, staff notation must be a visual Imagen, pitch rehearsal must be an auditory Imagen, and note names must be a auditory Logogen. Furthermore, the dual–coding by visualizing melodic contour along with pitch rehearsal would be considered the operation of a single system in Paivio's theory, because the melodic contour would be a visual Imagen and pitch rehearsal would be an auditory Imagen, and both codes would be together in a nonverbal system.

The results regarding explicit motor representations found in Experiment 7 provide some evidence supporting important views proposed in the levels–of–processing theory. The greater the degree of challenge in terms of melody length, ISI duration and interfering stimuli, the greater the efforts of the subjects to reduce the latency and to make elaborative and rapid finger movements. This finding supports the view that cognitive (conscious) effort is used for encoding (e.g., Battig, 1979 ; Jacoby & Craik, 1979 ; Lockhart, Craik, & Jacoby, 1976). It was also found that the finger movements gradually became more appropriate, accurate and rapid during the retention interval. This finding supports the view of a *continuum* of rehearsal operations running from the minimal processing necessary to repeat a word continuously to various types of elaborative processing involving either further enrichment of one item or associative linkage of several items (Craik, 1979), and the view that the encoding process gradually spreads and becomes richer and more elaborative (e.g., Anderson & Reder, 1979 ; Craik & Tulving, 1975 ; Lockhart, Craik, & Jacoby, 1976). Thus, the more the subjects rehearsed elaborative tapping, the greater the recognition performance became (e.g., Rundus, 1971, 1977, 1980 ; Rundus & Atkinson, 1970 ; Rundus, Loftus, & Atkinson, 1970).

Comparison of the motor rehearsal rate with the verbal and visual rehearsal rates revealed that the former is somewhat faster than the verbal rate, followed by the visual rate (e.g., Landauer, 1962 ; Norman, 1976 ; Rayner, 1978 ; Sternberg, Monsell, Knoll, & Wright, 1980 ; Thomas, Hill, Carroll, & Garcia, 1970 ; Vaughn 1983 ; Warren, 1969). When subjects are motivated to attempt to encode pitch information, the greater the number of times and faster the code is rehearsed, the deeper and more elaborately the code is processed. Motor encoding with finger movements might be relatively deep level of processing requiring much encoding effort at the first stage (it becomes gradually automatically), followed by verbal encoding with note names and, finally, visual encoding with melodic contour, which is a relatively shallow level of processing requiring less encoding effort (see Experiment 1).

Baddeley (1986, 1990) proposed the working memory model. The working memory system consists of three components, the most important of which is the central executive. The articulatory loop and the visuo-spatial sketch pad are subordinate systems.

The articulatory loop, in which verbal rehearsal occurs, consists of a passive phonological store and an articulatory control process. Originally a passive phonological store is directly concerned with speech perception. The neuropsychological evidence indicates that it is an auditory short–term store in which the processing is at a phonological level, and meaningful sounds, continuous speech and visually presented materials are not processed. An articulatory suppression task has little effect on the storage of phonological material. An articulatory control process is linked to speech production. An articulatory suppression task effectively prevents the use of "inner voice" or subvocal rehearsal.

Based on these properties of the phonological store and the articulatory control process, the short–term store of pitch information with acoustic level (auditory imagery), in which 'raw' pitch is mentally rehearsed, might be analogous in function to the phonological store, and the verbal rehearsal of note names (at accurate pitches) might be analogous to subvocal "singing" in the articulatory control process. Logie and Edworthy (1986) presented a finding that memory for tone sequences involves both mechanisms ; a phonological store and articulatory control process.

The visuo–spatial sketch pad specialized for spatial and/or visual coding was defined by Baddeley (1986) as "a system especially well adapted to the storage of spatial information, much as a pad of paper might be used by someone trying for example to work out a geometric puzzle." As found in this study, visualizing of pitch information was based on visual representations of melodic contour or staff notation , or visuo–spatial representations which were closely related to external or internal spatio–motor representations. The evidence that subjects tracked precisely their visuo–spatial image while visualizing pitch information was obtained by analyzing eye movements. These results are accounted for well by the visuo–spatial sketch pad concept.

CHAPTER 1

A REVIEW OF BACKGROUND CONCEPTS

Diversified Fields in the Psychology of Music

Music and Language

Historical Background in the Study of Memory

1. Multi–store model
2. Limitations to multi–store models
3. Levels–of–processing theory
4. Mental representations and dual–coding theory
5. Working memory model

Encoding and Representations

Outline of the Following Experiments

1. Preliminary experiment and its implications
2. Encoding strategy and distractor paradigm

Diversified Fields in the Psychology of Music

Since the mid–1970's, there has been remarkably progress in the study of music perception and cognition, reflecting the rapid pace of developments in cognitive psychology. Recent advances in computer technology have enabled investigators to generate various complex sound stimuli precisely. It has become possible to investigate such issues as auditory pattern analysis, the attentional mechanisms in music, and memory for musical information, and so on, with the stimuli control for strict experimentation (Deutsch, 1982a). Thus, the employment of generative approaches and theories has been progressively refined. At present, there are such diversified fields as pitch class (tone chroma) and octave similarity, absolute pitch, perception and cognition of pitch and harmony, processing of tonal and atonal pitch structures, rhythm and tempo, timing and dynamics, timbre of musical instruments, grouping and attentional mechanisms, music understanding and memory, music and space (representations and interactions) tuning, intonation and consonance, segregation and integration of tones, musical expectancy, singing voice, music performance (expression and style, technical and expressive musical skill acquisition, motor aspect), ability and training, musical education, prenatal audition, development and musical perception and cognition, musical creativity and cultural factors, ethnomusicology, cognitive musicology, brain function, psychophysiological or neuropsychological correlates of musical activities, aesthetic and affective response, emotional response to music, concert hall acoustics, and computer simulation of musical processing.

Music and Language

The similarity and discrepancy of music between speech has been noted in a number of contexts (see Repp, 1991 ; Sundberg, Nord, & Carlson, 1991). There are clear behavioral analogies between music and language (Gates, Bradshaw, & Nettleton, 1974). Both are composed of sequences of discrete sounds ; vocal production and auditory reception are primary in both. Both use arbitrary visual symbols to note salient aspects of the sound pattern. In addition, in both skills reading of the text at speed requires many years to develop, and it is reasonable to suppose that the music reader, like the language reader, increases his coding efficiency through attention to structures in the text.

However, the most radical difference between music and language or speech is that linguistics is directly related to the effect of the communicative message. Language is largely a stable referential symbolic system that represents our knowledge. Lexical items are symbols that exist not only by themselves in phonological terms, but also as representatives for their referents. These lexical–referential linkages are fixed and the combination rules are constant and universal for all languages (Aiello, 1994 ; Marin, 1982). Therefore, when one encounters unfamiliar language, one cannot store the words as not only semantic but also acoustic codes, and the words produce no visual images and arouse no feelings or emotion.

On the other hand, music consists of a few non–referential items that are combined according to the prevailing stylistic rules of harmony, melody, timbre, rhythm, or musical form (Marin, 1982). Since each of these factors does not correspond to a specific or fixed representation, we understand the musical message or meaning with our own mode, that is, there are individual differences. Music, even unfamiliar music, is likely to evoke some internal intermodal representations by arousing some feeling or emotion, and the pitch information can be encoded as some codes.

Music has a hierarchical structure which consists of several levels similar to those in language. Umemoto (1990) proposed four musical dimensions that have corresponding levels of perception and cognition. They are : (1) Music as sound (pitch, loudness, timbre, duration, and pitch class), which corresponds to discrimination or identification of tone ; (2) Music as an object of perception (melody, harmony, and rhythm), which corresponds to pattern or contour recognition, coding of tones in terms of a scale, and so on ; (3) the structure of music (theme and its development), which corresponds to comprehension of the structure ; and (4) the meaning or content of music (idea, title, and script), which corresponds to cognition and empathic understanding of the piece as a whole. Marin (1982) arranged musical deficiencies in a hierarchical order from a neuropsychological perspective, and distinguished disorders of the psychological, categorical, perceptual, lexico-symbolic, and programmative types.

Historical Background in the Study of Memory

1. Multi-store model

Prior to the recent remarkable progress in the study of music perception and cognition, cognitive psychology as an approach based on the information processing framework emerged in the 1950's. The information processing approach is partly based on a computer analogy. Information is handled in a sequence of stages ; in each stage a specified function is performed, and the information then proceeds to the next stage for a different kind of processing.

Several memory theorists have attempted to describe the basic architecture of the memory system. According to the Atkinson-Shiffrin model (Atkinson & Shiffrin, 1968, 1971), there are three types of memory stores ; the information from the environment is initially received by and held very briefly in the sensory stores, some is attended to and processed in the short–term store, and some is transferred to the long–term store.

Sensory memory holds information in a relatively raw, unprocessed form for an extremely short period. Most of the research which has been conducted on the sensory stores has concentrated on the visual and auditory modalities, which are the most important ones in our everyday lives. Neisser used the term iconic memory to describe visual sensory memory, or the brief persistence of visual impressions that "makes them briefly available for processing even after the stimulus has terminated" (Neisser, 1967, p.15). So far as the iconic store is concerned, the classic demonstration of its existence was provided by Sperling (1960). It is now generally accepted that information in iconic storage decays within approximately 500 msec.

There is also evidence for a transient store in the auditory modality. Neisser coined the term echoic memory as the auditory equivalent of iconic memory. Echoic memory is assumed to consist of relatively unprocessed auditory input, which persists after the sound disappears. Darwin, Turvey, and Crowder (1972) estimated that the duration of unattended auditory information in echoic storage is approximately 4 or 5 sec. Massaro (1970, 1972) used a masking technique, and concluded that echoic memory lasts about 250 msec. In another research, the duration was estimated to be about 100 msec (e.g., Deatherage & Evans, 1969 ; Efron, 1970).

Short–term memory (STM) refers to the small amount of information that we keep in an active state for a brief period. It includes the information that we are currently attending to, processing, and rehearsing to ourselves, However, the capacity of STM is extremely limited. A large proportion of materials is forgotten in STM after a few seconds' decay. Long–term memory (LTM) refers to a huge, relatively permanent kind of memory which has enormous capacity and contains a wide variety of information.

2. Limitations to multi-store models

The Atkinson–Shiffrin multi–store model served an important function historically. It was the first theory of memory which provided a systematic account of the structures and processes comprising the memory system. Although the conceptual distinctions between the three different kinds of memory stores still have conceptual utility, there are serious limitations to the multi–store model, in which both the STM and LTM are assumed to be unitary and to operate in a single, uniform fashion.

The short–term store was formerly held to accept only acoustically coded information. For example, Sperling (1960) had shown that the fading visual icon was rapidly translated into an auditory form. The most definitive experiment on this translation was conducted by Conrad (1964). Conrad presented letters visually in a memory span experiment, and found a strong tendency for people to make recall errors attributable to acoustical confusion. The results provide strong evidence that the short–term storage operates in an auditory mode. Similar finding was found by Wickelgren (1965). It was generally considered that only the long–term store could maintain semantic information. For example, Kintsh and Buschke (1969) found, in their experiment concerning the serial position, that the material in STM is coded in terms of its acoustic or sound characteristics, whereas that in LTM is coded in terms of its semantic or meaning–related characteristics.

However, not all storage in STM is in an acoustic format ; there is ample indication that verbal information is retained, at least to some extent, in visual form. Evidence that the short–term store holds visual codes has been found by several researchers (e.g., Kroll, 1972, 1975 ; Kroll & Kellicutt, 1972 ; Kroll, Parks, Parkinson, Bieber, & Johnson, 1970 ; Parkinson, 1972 ; Parkinson, Parks, & Kroll, 1971 ; Salzberg, Parks, Kroll, & Parkinson, 1971). Some experiments revealed that shadowing interfered less when the to–be–remembered letter was presented visually rather than acoustically. This suggests that visually presented letters are not recoded into pure acoustic information. The accuracy of response of deaf subjects is negatively affected by visual similarity of letter stimuli, resulting in recall errors attributable to visual confusion. This suggests that the deaf subjects retain letters in terms of their visual printed shapes (e.g., Conrad, 1972 ; Locke & Locke, 1971). That deaf subjects can retain verbal items in a code represented in their sign-language symbols has also been demonstrated (Bellugi, Klima, & Siple, 1975). The deaf have alternatives to the usual acoustic STM coding. In addition, there is evidence for the presence of semantic information in STM. Shulman (1970, 1972) found that some short-term confusions follow semantic patterns, indicating STM storage of semantic information. There is evidence for articulatory coding as well (Levy, 1971 ; Peterson & Johnson, 1971).

Another weakness of the approach in the multi-store model concerns the role of rehearsal, which is the cycling of information through memory. The model assumes that the major means by which information is transferred to LTM is via rehearsal in STM. In fact, while the amount of rehearsal is often relevant to LTM (Rundus & Atkinson, 1970), several researchers found that the amount of rehearsal in STM was not always related to the probability of recall from LTM (Craik & Watkins, 1973).

3. Levels-of-processing theory

Craik and Lockhart (1972) proposed a level-of-processing approach, in which there are a number of different levels of processing ranging from shallow (sensory or physical analysis of a stimulus) to deeper (more complex or semantic analysis). The by-product of all this analysis is a memory trace. If the stimulus is analyzed at a very shallow level, then that memory trace will be fragile and may be quickly forgotten. However, if the stimulus is analyzed at a very deep level, then that memory trace will be durable and will be remembered.

They also discussed the role of rehearsal. There are two different kinds of rehearsal. One is maintenance rehearsal, which simply repeats and maintains an item at a shallow acoustic level in STM, and cannot influence LTM. The other is elaborative rehearsal, which enriches and supplements the item with extra meaning at a deeper level, making it more likely to be stored in LTM ; that is, deeper levels of analysis produce a more elaborate, longer, and stronger memory trace (e.g., Craik & Lockhart, 1972 ; Craik & Tulving, 1975 ; Craik & Watkins, 1973).

The level–of–processing theorists emphasized encoding ; that is, how items are placed into memory. They have indicated that encoding process which takes place at the time of learning have a major impact on subsequent long–term memory and the process is important.

Furthermore, Craik (1979) proposed that when a stimulus is processed at a deep level for a long period, this stimulus is encoded distinctively. Distinctiveness describes the extent to which a stimulus is different from the other memory trace in the system. Eysenck (1979) also argued that long–term memory is affected by distinctiveness of processing as well as by the depth and elaboration of processing. In other words, memory traces which are distinctive or unique in some way will be more readily retrieved.

The major criticism of the levels–of–processing theory comes from Eysenck (1978), Jenkins (1974), and Baddeley (1978). The problem is caused by the lack of any independent or objective measure of processing depth. It seems intuitive that we can define physical encoding as "shallow" and semantic encoding as "deep". The future success of the levels–of–processing theory may depend upon finding a way to measure the critical feature, depth of processing.

4. Mental representations and dual–coding theory

Many of the proposed methods for improving memory involve the use of imagery, or mental pictorial representations for things that are not physically present. Although the hidden and covert nature of cognitive processes makes them extremely difficult to exam–

ine, the study of imagery, using new and more rigorous experimental methods, has begun to flourish again in recent years.

Representations may be divided into the two categories of the external representations, which we use in everyday life, and the internal or mental representations. The external representations may be divided into two broad classes of pictorial and linguistic classes, and the mental representations can be divided along similar lines into analogical and propositional representations. Although there are several forms of analogical representations (e.g., auditory, olfactory, tactile or kinetic images), the prime analog representation is a visual image. Analogical representations are nondiscrete, can present things implicitly, are organized by loose rules of combination, and are concrete in the sense that they are tied to a particular sense modality (modality-specific). Propositional representations are language-like representations which capture the ideational content of the mind, irrespective of the original modality in which that information was encountered. They are discrete, explicit, organized by strict rules, and abstract.

Many theorists claim that information is stored in analog code, which is a representation that closely resembles the physical object (e.g., Kosslyn, 1980 ; Cooper & Shepard, 1973 ; Shepard & Metzler, 1971). Some of the characteristics of mental images are the phenomenon that they can be mentally rotated and scanned, with preservation of relative size and shape, in ways analogous to such operations on physical objects. Several dissenters believe that we store information in terms of abstract descriptions of objects or propositions ; storage is verbal rather than visual or spatial (Pylyshyn, 1973).

In a study of how imagery influences memory, Paivio (1969) obtained the result that people recall more concrete words (e.g., apple, house) than abstract words (e.g., idea, truth), indicating that concrete words encourage imagery. Paivio (1978) explained the result using a dual–coding hypothesis in which there are two distinct systems for the representation and processing of information. A verbal coding system deals with linguistic information and stores it in an appropriate verbal form, and it is specialized for sequential processing to the serial nature of language. A separate nonverbal (or imagery) coding system carries out image–based processing and representations that correspond to concrete objects (Bower,1972a). Although the two systems are independent, they are connected to each other and can cooperate with each other. Evidence for separate verbal and non–verbal systems has been observed in patients suffering unilateral damage to the temporal lobes (Paivio, 1971).

The hypothesis maintains that superior memory for words with high imaginability arises because they are readily encoded in both the imaginal and verbal systems, whereas words with low imaginability are likely to be encoded only in the verbal system. Highly imaginable words are better remembered because they are represented in the memory as two types of representations rather than one. Paivio (1986) also proposed that both the verbal (Logogens : the concept comes from Morton's (1969) theories of word recognition) and nonverbal (Imagens) systems are further divided into sub–systems (i.e., visual, auditory, kinesthetic, gustatory, and olfactory) which process either verbal or nonverbal information in the different modalities (modality–specific).

5. Working memory model

When the multi–store model fell into disfavor, as described above, Baddeley and Hitch (1974) proposed the concept of the working memory. In essence, the unitary short–term store was replaced by a multi–component working memory system consisting of three components : a modality–free central executive which has limited capacity and has two subordinate components, an articulatory loop which holds information in a phonological (speech–based) form, and a visuo–spatial scratch pad (sketch pad) which is specialized for spatial and/or visual coding.

The articulatory loop is organized in a temporal and serial fashion, and its capacity is determined by temporal duration. Baddeley (1986) proposed a revised version of the working memory model. He drew a distinction between a passive phonological store which is directly concerned with speech perception and an articulatory control process which is linked to speech production. The visuo-spatial sketch pad is defined as a system especially well adapted to storage of spatial information (Baddeley, 1986), as Baddeley and Lieberman (1980) found that spatial coding was more important than visual coding in a variety of tasks. While the revised working memory model is probably a refinement of the original model, it is unfortunate that there has been so little clarification of the role played by the central executive. It is claimed that the central processor is modality-free and used in numerous processing operations, but its precise functioning still remains unclear.

Encoding and Representations

The purpose of this study was to investigate the encoding strategies for pitch information of short melodies. "Encoding" refers to the establishment of codes which are more likely to be stored in relatively permanent long-term storage. The encoding processes are often assigned in the short-term memory.

Several authors (e.g., Bower, 1967 ; Tulving & Watkins, 1975) have suggested that the memory trace can be described in terms of its component attributes. This viewpoint is quite compatible with the notion of encoding elaboration. The trace may be considered the record of encoding operations carried out on the input ; the function of these operations is to analyze and specify the attributes of the stimulus. An encoded unit is unitized or integrated on the basis of past experience, just as the target stimulus fits naturally into compatible context at encoding. An integrated or congruous encoding thus yields better memory performance, first, because a more elaborate trace is laid down and, second because richer encoding implies greater compatibility with the structure, rules, and organization of semantic memory. This structure, in turn, is drawn upon to facilitate retrieval processes (Craik & Tulving, 1975).

To acquire effective encoding processes for verbal materials, there are many types of strategies, such as a elaborative rehearsal, use of imagery, enactment of verbal instructions, utilizing semantic relations between list items to integrate them into organized memory units, and using various mnemonic devices (including chunking, grouping, organization, mediators, first-letter technique, narrative technique, substitution method, rhymes, and external memory aids, and so on).

A variety of mental representation models have been proposed for pitch structures. These models fall into one of three classes of models (see Dowling, 1991 ; Krumhansl, 1990, 1991 ; Hubbard & Stoeckig, 1992 ; West, Howell, & Cross, 1991) ; (a) psychoacoustical, (b) rule-based, and (c) schematic / connectionist approaches. Psychoacoustical models hypothesized that pitch could be represented by a single dimension (e.g., Stevens, Volkmann, & Newman, 1937). Rule-based models exploit some of the outward similarities between music and language. In language, there are a limited number of letters that can be combined to form valid words, and words that can be combined to form grammatical sentences. In music, there a limited number of tones (12 tones in a chromatic scale) that can be combined to form melodies and accompanying harmonies. Deutsch (1980) and Deutsch and Feroe (1981) presented a hierarchical model represented by pitch alphabets. Lerdahl and Jackendoff (1983) developed an extensive generative grammar of music founded on four sets of rules ; grouping structure, metrical structure, time-span reduction, and prolongation reduction. In schematic models, pitch is represented by some sort of music schema which consists of at least three, or more dimensions ; Longuet-Higgins' key regions model (1978), Shepard's double–helix model (1982a,b), Rectangular representation of the multidimensional scaling analysis of interkey distances (Krumhansl & Kessler, 1982). Lerdahl's model (1988) is designed to capture essentially the same feature of musical pitch as Shepard's (1982a,b) topological model, but is more symmetry on pitch structure. Connectionist model of the representation of harmonic relationship was introduced by Bharucha (1987) and Bharucha and Stoeckig (1986, 1987). In this model musical chords are represented by nodes in a network.

There are a lot of models of the mental representations for pitch structures as described above, indicating that any given music stimulus is likely to have multiple representations of pitch, and that, therefore, people would employ multiple coding strategy in auditory modality for pitch sequences (West, Howell, & Cross, 1991). Shepard (1982a) suggests that for musically sophisticated listeners, pitch should be encoded by three dimensions ; height, position in the circle of fifths, and chroma. These three dimensions are all in auditory modality (though they are visualized as a figure on a sheet : helical configuration of tones, with pitch height as the vertical dimension and the chroma circle as the projection onto the horizontal plane). However, when pitch sequence is presented auditorily, do subjects actually encode the pitches with the mental representation of the helical configuration with the chroma circle ? The helical configuration with the chroma circle is a sophisticated theoretical and conceptual representation for pitch structure, but is not likely to be auditory (or visual) imagery perceived from pitch information. Even though a rule–based or schematic / connectionist model might offer us a complete description of how musical representation might be accomplished, it is not clear that such an understanding or such a representation would necessarily include the subjective elements of imagery so commonly reported (Hubbard & Stoeckig, 1992). Furthermore, it is possible that auditory pitch information is encoded not only in auditory modality, but also in visual or kinesthetic modality, namely that multimodal representations based on intermodal coordination are employed to encode auditory pitch information.

Therefore, in the present study, encoding strategies for pitch information in short melodies were investigated, in terms of not rule– or schema–based representations but imagery–based multimodal representations.

For a time in the 1970' and early 1980', the imagery–propositional debate was very controversial. During the period, Posner (1973) presented the conclusions concerning codes in memory that (a) there are at least three types of codes, visual, verbal, and motor (Bower, 1972b) ; (b) each code endures, and is not a transient residual of stimulation ; (c) people differ in their propensity to use each type of code ; (d) these codes are parts of separate memory systems that can be examined in isolation in the laboratory. In this study all four of the points were experimentally demonstrated.

Outline of the Following Experiments

1. Preliminary experiment and its implications

The purpose of this study was to investigate encoding strategies for pitch information of short melodies. First of all, a simple preliminary experiment was carried out. Subjects were instructed to make recognition judgments as to whether an 8–tone standard melody and an 8–tone comparison melody were the same or different in pitch. The subjects were then required to answer the following question ; when you memorized and retained the melodies, what types of strategies did you use to encode the pitch information ?

The preliminary questionnaire data indicated that the subjects used one or more of the following several strategies : (a) a verbal encoding strategy, in which each pitch in a melody was labeled with the name of a musical note, and this code was rehearsed and stored in memory ; (b) a sensory (auditory) encoding strategy, in which pitches in a melody were retained in memory as auditory information ; that is, by singing, whistling, humming, mental rehearsal of pitches, and so on ; (c) a visualizing strategy, in which pitches were visualized in their image, as a melodic contour, on a keyboard, or on a staff notation ; or (d) a motor encoding strategy, in which an auditory melody was encoded by the movement of the fingers as if playing the piano. Some subjects reported that they used two or three strategies simultaneously.

The questionnaire data obtained in this simple experiment presented several important implications, which were investigated in the subsequent experiments. There may be two rehearsal modes, that is, pitch rehearsal which is direct and relatively unprocessed pitch representation and verbal rehearsal of musical note names (Experiment 1). It is possible that visual imagery, which is analogous to auditory information, is used to encode pitch information (Experiment 4). These pitch encoding strategies obtained in the preliminary experiment were not entirely independent of each other, and were sometimes inter–related. This observation implies that a dual– or a triple–coding strategy may be used to memorize or retain pitches (Experiment 6). It is possible that the employment of finger movements analogous to playing the piano may be an effective strategy for piano players to encode pitches, because this external motor representation is likely to be resistant to challenging situations (Experiment 7). On the keyboard of a piano, higher pitch keys are placed on the right and lower pitch keys are placed on the left. Piano players may have an image of a spatial configuration in which the right direction is compatible with higher pitch, and the left with lower pitch. It is possible that there are specific spatial directions for pitch corresponding to spatial motor images during playing of an instrument (Experiment 8).

In Experiment 2, the ability to detect a 50–cent deviated pitch from an equal–temperament scale out of a melody was investigated. This ability is closely related to the ability for verbal encoding of pitch information. In Experiment 3, the relationship between the ability for verbal encoding of pitch information and cerebral hemispheric dominance was investi–

gated. Experiments 5 and 9 were designed to investigate whether subjects show eye movements corresponding to tracking of their visuo-spatial images while encoding pitch information.

2. Encoding strategy and distractor paradigm

The methodology used in the following experiments was influenced by that of Sloboda's experiments (1976). One means of attempting to investigate the nature of the encoding processes is to interfere with them by causing concurrent stimulation or activity which engages in similar processing mechanisms (distractor paradigm).

The purpose of the distractor task is to prevent rehearsal. A distractor task leads to increased incidence of forgetting, if it involves processing of stimuli that are similar to the to-be-remembered materials (e.g., Corman & Wickens, 1968 ; Wickelgren, 1965). In the experiments of Reitman (1971, 1974) and Shiffrin (1973), it was found that distractor tasks involving verbal skills were more likely to disrupt retention of verbal forms than were non-verbal distractor tasks such as signal detection. Furthermore, the distractor task is much more effective if it presented in the same modality as the to-be-remembered materials (e.g., Proctor & Fagnani, 1978). The distractor paradigm is also accounted for by a mechanism of release from proactive inhibition (e.g., Keppel & Underwood, 1962 ; Wickens, 1972 ; Wickens, Born, & Allen, 1963).

The purpose of Sloboda's experiment (1976) was to investigate a code in a visual memory task. The visual stimulus was six notes on a staff, and was presented in a two-field tachistoscope. In each trial, the visual stimulus was presented for 2.0 sec, and the subject was then required to recall the notes on a staff. Three types of interfering material (speech, tonal music, and atonal music) were prepared for auditory presentation, and presented as the interference condition. It was hypothesized that if musicians coded the visual notes as pitches, then they would suffer greater interference from the tonal or atonal musical input in the auditory modality than from the speech input, whereas if they coded the visual notes as note names, then they would suffer greater interference from the speech input. The results were interpreted to indicate that either the musicians did not encode visually presented notes by their note names or pitches, or that they can carry on concurrent activities with the same code simultaneously.

The purpose of the present study was to investigate the types of strategies used to encode the pitch information in short-term memory. Therefore, comparing Sloboda's experiment, the stimuli were presented auditorily instead of visually, and the subjects were required to produce recognition responses instead of recall responses. Although a distractor paradigm was employed in Sloboda's experiment, it is possible that the distractors did not interfere with the subjects' coding processes effectively, because, the distoractors may not have been similar to their predictable coding strategies, or the modality of distractors were different from that of the to-be-remembered stimuli.

Therefore, in the present experiments, interfering stimuli which more effectively interfere with the subjects' coding processes than those used by Sloboda were introduced. Tones or note names, whose rate was the same as that in the standard melody, were interpolated auditorily during the retention interval (Experiments 1, 6 and 7), instead of the continuous music or speech used as background interference by Sloboda.

In Experiments 1 and 4 to 9, to investigate the type of encoding strategies, several kinds of distractors were prepared and interpolated during the retention interval between the standard and comparison stimulus. Distractors interfere with the operation of the codes which are used to acquire the pitch information of the standard stimulus and are maintained or rehearsed. From the disruptive effects of distractors on the recognition performance, the type of encoding strategy employed during retention interval can be inferred.

CHAPTER 2

EXPERIMENTAL TESTS OF VERBAL ENCODING STRATEGY FOR PITCH

EXPERIMENT 1

Verbal Encoding Strategy and Pitch Rehearsal Strategy

Method

Results and Discussion

EXPERIMENT 2

Detection of Deviated Pitch out of Tonal and Atonal Melodies

EXPERIMENT 2–1 *Method*

Results and Discussion

EXPERIMENT 2–2 *Method and Results*

EXPERIMENT 3

Cerebral Hemispheric Dominance for Various Types of Melodies

EXPERIMENT 3–1 **For Western melodies**

Method

Results and Discussion

EXPERIMENT 3–2 **For Japanese melodies**

Method

Results and Discussion

EXPERIMENT 3–3 **For tonal and atonal melodies**

Method

Results and Discussion

EXPERIMENT 3–4 **For tonal and atonal melodies along with note names**

Method

Results and Discussion

EXPERIMENT 1

Verbal Encoding Strategy and Pitch Rehearsal Strategy

In Deutsch's experiment (1970), subjects were required to judge whether two tones, separated by a 5–sec interval, were the same or different in pitch. The tones were of the same pitch in half the trials, and differed by a semitone in the other half. Results showed that when six extra tones were interpolated during the 5–sec retention interval, the error rate was 32.3 %, even though the subjects were required to ignore the interpolated tones. In contrast, when six spoken numbers were interpolated instead of tones, the error rate was only 5.6 %, even though the subjects were required to recall the numbers. Similar results have been obtained in other studies ; whereas only a minimal decrement in pitch recognition was found after a retention interval filled with a sentence, noise or a series of numbers, considerable disruption was found when several tones were interpolated (e.g., Massaro, 1970 ; Wickelgren, 1966).

On the other hand, Siegel (1974) found that, if the subjects were able to encode pitch verbally ; i.e., to label tones with the musical note names and to store them in memory, pitch recognition did not drop significantly even after a 15–sec delay filled with interfering tones. Verbal encoding thus aided retention over a longer time span.

These results indicate that pitch information is subject to interference caused specifically by other tones, and that only when some effective strategies can be employed to encode pitch information will it be retained in memory. Therefore, the purpose of Experiment 1, with distractor paradigm (Umemoto, 1984 ; Umemoto, Takeda, Arimoto, & Kondo, 1976), was to investigate whether subjects could employ verbal encoding as an effective strategy to encode pitches of tonal and atonal melodies (Mikumo, 1989a,b, 1990b, 1992d), instead of only one tone, as in the experiments described above.

Method

Subjects

Twenty-six female subjects (average age, 21.3 years ; age range, 20–23 years) who were undergraduates majoring in music constituted the highly musically trained group (Group H). Each had had at least 12 years of formal musical training, with an average of 15.7 years (range, 12–19 years).

Twenty-six female subjects (average age, 20.6 years ; age range, 18–23 years) who were undergraduates majoring in literature or science constituted the less well musically trained group (Group L). None were currently playing a musical instrument, and their musical experience on an instrument averaged 3.8 years (range, 0–5 years).

Materials

Each trial involved a standard series followed by an interfering stimulus and then by a comparison series. Both the standard and the comparison series consisted of six different tones taken from an equal–temperament scale, which ranged from G_3(196.0 Hz) to G_5(784.0 Hz). The duration of each tone was 1.0 sec (900msec with 100–msec silence). In this experiment there were 64 trials. Half of the them (32) were tonal melodies in a major key and were high in tonal melodic structure, according to conventional Western rules. The other half(32) were atonal melodies, which were low in tonal melodic structure(1).

Half of the comparison series were exactly the same as the standard series, and the other half were divided into four types that were different from the standard. Transposition (T) was obtained by transposing the standard series by two semitones (higher or lower), preserving both the exact pitch intervals and the contour[2]. Contour–preserving comparison (C) was obtained by changing one of the pitches by two semitones (higher or lower), preserving the contour of the standard, so that the exact pitch intervals were not preserved. Exchanging comparison (E) was obtained by exchanging the order of two successive pitches of the standard, so that both the contour and the pitch intervals were a little different from those of the standard. Retrograde comparison (R) was obtained by reversing the order of pitches of the standard, so that the contour and the pitch intervals were not completely preserved.

Four types of interfering stimuli were interpolated during the retention interval between the standard and the comparison series. "Pause" (P) was a blank retention interval. "Interfering Melody" (IM) was composed of 12 tones (range, G_3–G_5), having either a tonal or an atonal structure. "Series of Nonsense Syllables" (NS) consisted of a meaningless series of 12 speech sounds. "Series of musical Note Names" (NN) consisted of a series of 12 words chosen randomly from seven note names (Do, Re, Mi, Fa, Sol, La, Si). Both the series of nonsense syllables and the series of note names were delivered in a monotonous female voice. Each of four interference types consisted of 16 stimuli, having a total duration of 12 sec (Figure 1–1).

The tones were generated by an NEC PC–8801 MK2 computer, and were recorded on tape and played over high–quality sound reproduction equipment. All tones were adjusted to be equal in loudness (approximately 50dB SPL), and were presented via two loudspeakers.

Procedure

The subjects were instructed that this was an experiment on memory for melodies, and that, on each of the 64 trials, they would first hear a warning signal, then a first melody, which was followed by an interfering stimulus, and then a second melody. The subject's task was to judge whether the two melodies were the same or different in pitch, indicate their judgments by writing "S"(Same) or "D"(Different) on an answer sheet, and rate their decision on a 5-point confidence scale with responses of "very sure yes (or no)", "fairly sure yes (or no)", "unsure yes (or no)", "fairly unsure yes (or no)", "very unsure yes (or no)".

As described above, four types of interfering stimuli were prepared. The subjects were required that, when the interfering stimulus was a pause or a melody, they should listen to it, and when the interfering stimulus was a series of nonsense syllables or of note names, they should shadow the sounds one by one. The experimenter emphasized that the subjects were to respond "different", even to exact transpositions of the first melody, and that they should not employ a tapping strategy analogous to playing the piano. The subjects were given feedback on four practice trials for each of the interfering stimulus types. The intervals between the standard stimulus and the first interfering sound, and between the last interfering sound and the comparison stimulus, were each 1.0 sec, and the next trial began 7 sec after the comparison stimulus.

In this experiment, there were 32 tonal and 32 atonal trials. In both types of melodies, 8 trials were run in each type of interfering condition, in half of which (4 trials), the comparison series were exactly the same as the standard series, and in the other half (4 trials), the comparison series were different from the standard series (T, C, E, R). The four types of interference conditions were mixed, and the order of presentation of all 64 trials was randomized.

Figure 1–1.

Examples of standard, comparison, and interference conditions used in this experiment. Filled notes in comparison stimulus sets indicate that their pitches were different from those located at the corresponding serial positions in the standard stimulus sets.

Notes.

Interference conditions

(P : Pause, IM : Interfering Melody, NS : series of Nonsense Syllables, NN : series of musical Note Names).

Comparison stimuli

(T : Transposition, C : Contour–preserving, E : Exchanging, R : Retrograde).

Results and Discussion

In this experiment three response measures were calculated : the recognition probability for each type of interference condition (Figure 1–2), the false–alarm rate for each type of comparison (Figure 1–3), and the receiver–operating characteristic (ROC) curve (Figure 1–4).

First, the recognition probability data (hit rate minus false–alarm rate [Woodworth & Schlosberg, 1954]) were analyzed in a three–way analysis of variance [2 Groups X 2 Melody Types X 4 Interference Conditions], with repeated measures on the second and third factors. There were significant main effects of Group, Melody Type, and Interference, [$F(1,50)=150.8, p<.001$; $F(1,50)=64.9, p<.001$; $F(3,150)=7.06, p<.001$], and there was a marginal interaction of Group X Melody Type X Interference [$F(3,150)=2.47, p=.063$]. These results are shown in Figure 1–2, in which the main effect of Group indicates that the performance of Group H was consistently superior to that of Group L, and the main effect of Melody Type indicates that both Groups H and L performed better with the tonal than with the atonal melodies. The marginal interaction among the three factors indicates that there might have been differences in the disruptive effects of the interference conditions in the performance of the two groups.

As shown in Figure 1–2, recognition performance of Group H for the tonal melodies was significantly (by Newman–Keuls Method) more disrupted by shadowing the series of note names than by the Pause ($p<.01$), or by shadowing the series of nonsense syllables ($p<.05$). This finding suggests that Group H used a verbal encoding strategy, in which each pitch in a tonal melody was labeled with a note name, and that these codes (note names) were rehearsed. For atonal melodies, recognition performance of Group H was

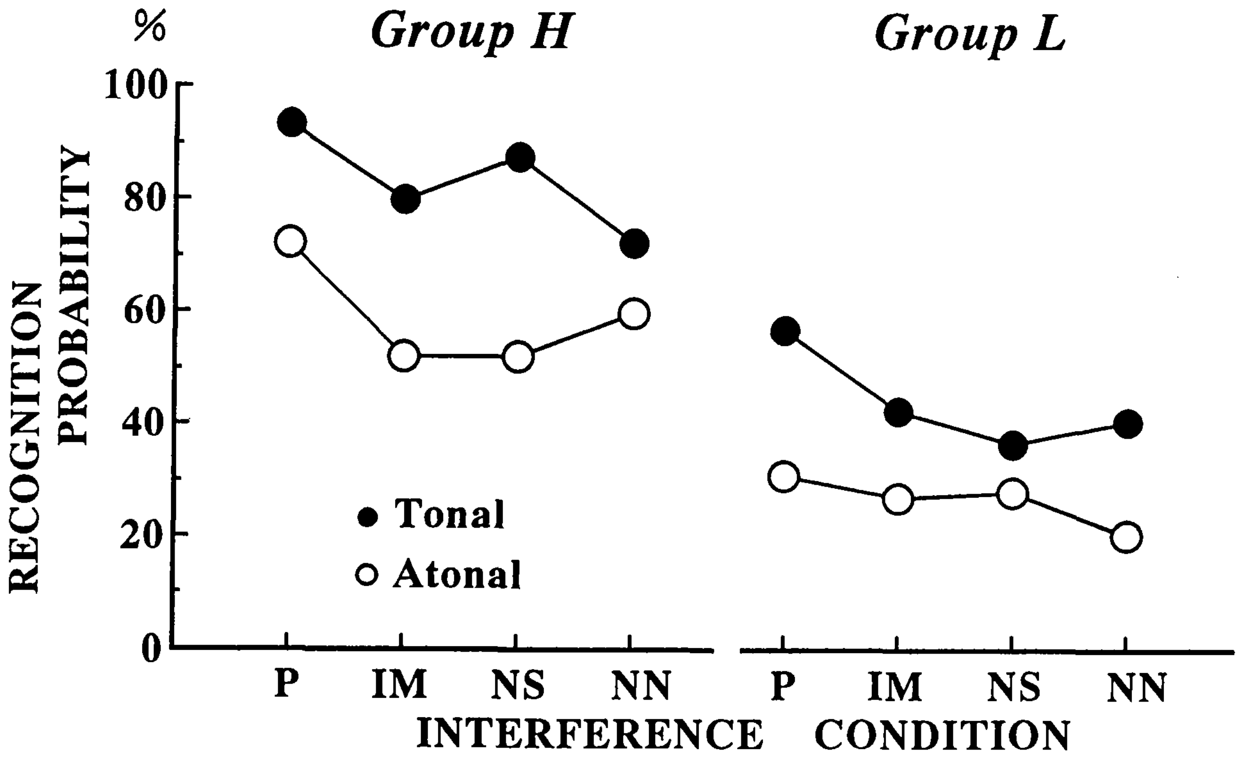


Figure 1-2.

Mean probability of correct recognition (hit rate minus false-alarm rate) in the four interference conditions (P, IM, NS, NN) for both tonal and atonal melodies.

Notes.

P : Pause, IM : Interfering Melody, NS : series of Nonsense Syllables, NN : series of Note Names.

significantly more disrupted by listening to the interfering melody than by Pause ($p<.05$). This finding suggests that the subjects used a pitch rehearsal strategy, in which they attempted to retain the pitches in the melody as acoustic information ; that is, by humming, whistling, singing, mental rehearsal of pitches and so on. These subjects reported that a verbal encoding strategy would be ineffectual for an atonal melody. Only very specific information about pitch is extracted from atonal melodies encoded at the material level of tones. The pitches and intervals of atonal melodies seem especially hard to learn (cf. Dowling, 1982, 1988).

On the other hand, the recognition performance of Group L for both the tonal and the atonal melodies was significantly worse than that of Group H even in the Pause condition, and there were no significant differences in disruptive effects for the different types of interference conditions. This finding suggests that these subjects could not use any effective strategies for pitch encoding of the melodies.

Second, the false-alarm data were analyzed in a three-way analysis of variance [2 Groups X 2 Melody Types X 4 Comparison Types], with repeated measures on the second and third factors. There were significant main effects of Group and Comparison Type, [$F(1,50)=36.3, p<.001$; $F(3,150)=29.3, p<.001$], and there were significant interactions of Group X Comparison Type, and Group X Melody Type X Comparison Type, [$F(3,150)=7.90, p<.001$; $F(3,150)=3.07, p<.05$], but there was no significant main effect of Melody Type. These results are shown in Figure 1–3, in which the interaction of Group X Comparison Type is shown. For Group H, significant differences (by Newman–Keuls Method) in false-alarm rate were found between Contour-preserving and Exchanging comparisons, Contour-preserving and Retrograde comparisons, Exchanging and Retrograde comparisons, and Contour-preserving and Transposition comparisons ($p<.01$). For Group L, significant differences were found between Transposition and Exchanging comparisons, Contour–preserving and Exchanging comparisons (p<.05), Transposition and Retrograde comparisons, Contour–preserving and Retrograde comparisons, and Exchanging and Retrograde comparisons (p<.01). The false–alarm rate of the former of each pair was higher than that of the latter. Namely, Group L had many more false alarms in Transposition than had Group H.

These findings suggest that the subjects in Group L were unable to discriminate a difference between the standard and the transposition, even though they were asked to reject the exact transposition of the standard, that is, pitch or key was not effective cue. In other word, they could use only the contour as a cue for melody recognition.

The subjects in Group H were able to use not only contour but also pitch intervals as effective cues for melody recognition. Cuddy, Cohen, and Miller (1979) have also shown that if the sequence to be transposed is embedded in a tonal context, it is easier to recognize than when in a non–tonal context. In the present experiment, stimuli would be somewhat easy for Group H, so that even with atonal melodies the false–alarm rate was rather low in transposition.

These interpretations suggest that contour and pitch intervals are independently handled as different features (Edworthy, 1985), and moreover, suggest that contour recognition is dominant over pitch intervals recognition. These results are consistent with the findings of Deutsch (1977), Dowling (1971, 1972, 1978), Dowling and Fujitani (1971), Idson and Massaro (1978), Kallman and Massaro (1979), and Sloboda (1978). They also suggest that whereas the subjects in Group L tended to listen global, the subjects in Group H were able to listen more analytically. This is consistent with the findings of Bever and Chiarello (1974) and Mikumo (1987a, 1988).

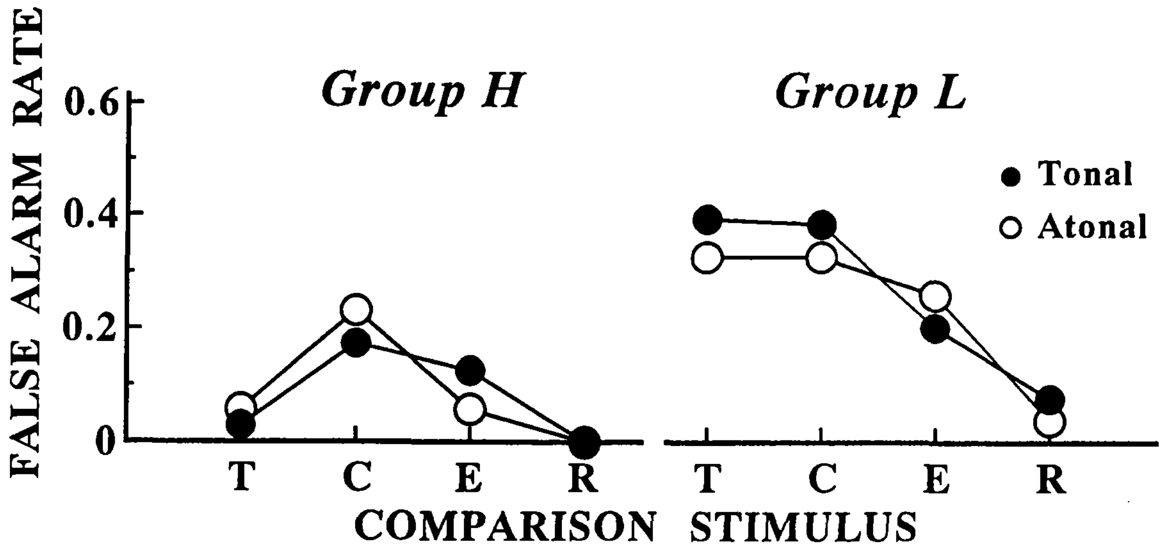


Figure 1–3.

Mean false–alarm rate for four comparison melodies (T, C, E, R) for both tonal and atonal melodies.

Notes.

T : Transposition, C : Contour–preserving, E : Exchanging, R : Retrograde.

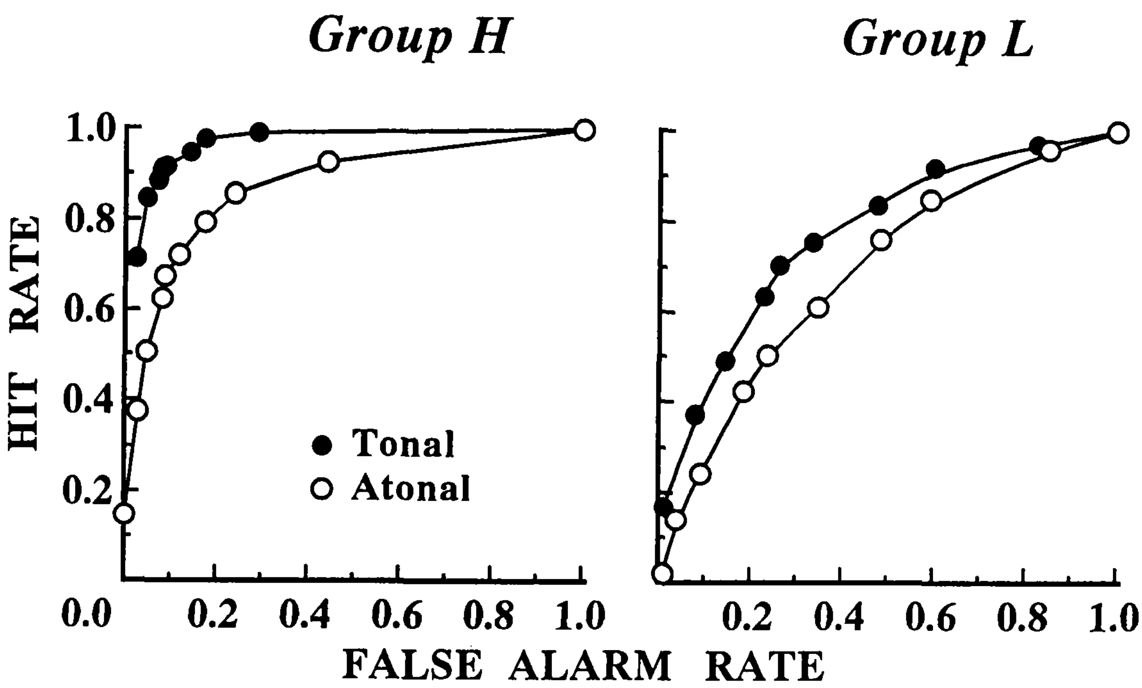


Figure 1–4.

Receiver Operating Characteristic (ROC) curve for recognition memory for tonal and atonal melodies.

Third, Figure 1–4 shows the ROC curve calculated by using a 5–point confidence scale, in which the "Same" judgment and the "Different" judgment each had five levels of response. The data were thus divided into 10 categories, plotting the hit rate of Same–Different decisions for each of the false–alarm rates. In addition, the discriminability indices (***d'***) were calculated for two types of melodies for the two groups (Group H: tonal melodies, d'=2.44 ; atonal melodies, d'=1.74. / Group L: tonal melodies, d'=1.12 ; atonal melodies, d'=0.64). It was evident from the d' and the ROC curve data that the discrimination in the Same–Different decisions of the Group H was consistently better than that of Group L, and both Groups H and L had more discrimination for tonal than for atonal melodies. Listeners were better able to detect alternations in tonal than in atonal melodies (cf. Dewar, Cuddy, & Mewhort, 1977).

It may be concluded that the less well musically trained subjects tended to listen globally to the musical series on the basis of the total configuration or Gestalt (i.e., melodic contour : cf. Dowling, 1994). Although they performed somewhat better with tonal than with atonal melodies, they were unable to use any effective strategy to encode pitches of these melodies.

On the other hand, the highly musically trained subjects could listen analytically to the musical series, with considerable attention to the internal relationships among their components (i.e., pitch intervals). When the series had a tonal structure, the subjects were able to use a verbal encoding strategy. Pitch information is mapped onto a relatively small set of highly overlearned alphabets. Although these differ from one culture to another, the use of such alphabets appears to occur cross–culturally (Deutsch, 1982c). Tonal sequences are coded and retained as hierarchies of structures, each of which is associated with a given pitch alphabet (Deutsch & Feroe, 1981).

Siegel (1974) proposed that possessor of absolute pitch attempted to use a verbal encoding strategy when the pitches belonged to different categories of note name. However, the results obtained in Experiment 1 showed that when the tones were serialized in an atonal structure, it was difficult even for the highly trained subjects to use a verbal encoding strategy even when each pitch in the melody belonged to a different category of note name.

In this experiment, one of four interference conditions was interpolated during the retention interval, and all of the results were interpreted relative to the control condition P (blank-interval). This is based on the view that forgetting in short-term memory comes about through both a decay in time and interference from the presentation of other material (e.g., Lindsay & Norman, 1977 ; Massaro, 1970 ; Norman & Bobrow, 1975 ; Wickelgren, 1967, 1973). In Experiments 4, 6, and 7 as well, a blank-interval control condition was interpolated during a retention interval to remove the disruptive effects of the time decay and to clear the disruptive effect of the interference.

In the results of this experiment, a ceiling effect was found in Group H with tonal melody, indicating that remembering a 6-tone melody was rather easy for the highly trained subjects. To investigate encoding strategies for pitches of melodies, which involve various information regarding contours and pitch intervals, an appropriate melody length would be more than 8 tones (cf. Edworthy, 1985 ; Mikumo, 1987a, 1988, Mikumo & Umemoto, 1988 ; Umemoto & Mikumo, 1988, 1989). In the following experiments except for Experiment 2, 8-tone melodies were employed as stimuli.

In the second analysis, it was found that, contour is easier to extract from melodies than is pitch or pitch interval information but contour is not easier to retain (cf. Dowling, 1982). Therefore, according to the level-of-processing theory (Craik & Lockhart, 1972), use of contour would reveal processing at shallow level and use of pitch intervals that at deeper level in terms of melody recognition. Moreover, when subjects perceive an auditory contour, a visual contour would be perceived at the same time ; that is, they would visualize an auditory contour as a visual contour. This point is discussed in more detail in Experiment 4.

Subjects in Group H reported that they employed verbal encoding strategy ; in most cases the verbal codes were accompanied by corresponding pitch codes. It is possible that two codes are employed at the same time to encode pitches. This point is discussed in more detail in Experiment 6.

To provide appropriate musical note names for the particular pitches (the encoding of pitch information as verbal labels) indicates that pitches separated by more than a semitone are categorized into different verbal labels (note names). The possession of the verbal labels increases the likelihood of the categorical information being extracted and stored (Sloboda, 1985). The employment of verbal encoding strategy for pitch information would be closely related to categorical perception for musical intervals. This point was investigated in Experiment 2.

****

(1) These stimuli were chosen from 100 original melodic series, each composed of six tones, by 20 university students using the scales of "Tonality-feeling" described in Hoshino (1985), and Hoshino and Abe (1981, 1984). Abe(1987), Abe and Hoshino (1985), and Erickson (1984) investigated tonal centering. Boltz (1989) investigated the effects of tonal relationships on melodic completion.

(2) Contour refers to an ordinal representation of the intervals of a melody, which indicates whether adjacent notes are higher or lower than one another (Jones, 1976).

EXPERIMENT 2

Detection of Deviated Pitch out of Tonal and Atonal Melodies

In Experiment 1, it was found that highly musically trained subjects were able to encode pitches in tonal structure as verbal labels (note names). The ability to assign different verbal labels to the pitches, which were separated by more than one semitone, was closely related to categorical perception.

Siegel (1974) proposed that possessors of absolute pitch had two distinct modes of pitch processing : a sensory trace mode, in which they attempted to retain some sensory information from tones for a brief time in memory, and a verbal mode, in which they attempted to label tones with the names of musical notes and to store these in memory. In Siegel's experiment, the subjects were presented with two tones, separated by a 5–sec interval, and they were required to judge whether the second tone was higher or lower in pitch than the first. The results indicated the operation of sensory coding when the two pitches were only one–tenth of a tone apart, since verbal coding would have provided the same name for both pitches. When the two pitches were separated by three–quarters of a tone, they belonged to different categories of note names and were differentially labeled ; therefore, the verbal labeling strategy was useful to subjects with absolute pitch.

Categorical perception of pitch intervals was demonstrated by Locke and Kellar (1973). They presented subjects with triads (chords) in which the tuning of the middle pitch was systematically varied from major (C_5 sharp=554 Hz) to minor (C_5 natural=523 Hz), passing through several intermediate pitches. The lower and upper pitches were always 440Hz(A_4) and 659 Hz(E_5). The results indicated that almost all chords with middle pitch above 546

Hz were heard as A major, while almost all chords with middle pitch below 540 Hz were heard as A minor. The evidence suggests a category boundary at about 542 Hz. Whereas musicians showed a clear peak in discrimination at the category boundary, non–musicians did not show such clear evidence of categorical perception.

A number of experimenters have obtained category–scaling identification functions for pitch intervals spaced at increments of 10 to 20 cents, over ranges of 2 to 5 semitones, where the labels are the relevant intervals in the chromatic scale (e.g., Burns & Ward, 1974, 1978, 1982 ; Rakowski, 1976 ; Siegel & Siegel, 1977a,b ; Umemoto & Tokumaru, 1990). In musically trained subjects, the form of the identification functions also shows very sharp category boundaries (Burns & Ward, 1978) and high test–retest reliability, whereas in musically untrained subjects, it shows large category overlap and poor test–retest reliability (Siegel & Siegel, 1977a).

These results indicate that, the closer a pitch is to one in the equal–temperament chromatic scale, the more the pitch is assimilated to the category on the scale. Up to 50 cents (one half semitone), the farther a pitch is from one in the equal–temperament chromatic scale, the more the pitch is dissimilated from the category on the scale, and beyond 50 cents up to 100 cents, the pitch is increasing assimilated into the next category on the scale. The category boundaries are very sharp in highly trained subjects.

In Experiment 2–1, subjects were required to identify a deviated tone in a melody, in which one tone was deviated by 50 cents either higher or lower in pitch from that in the equal–temperament scale. Fifty cents is one halftone between semitones, and a 50–cent deviated tone is in other words just a middle pitch between the two correct pitches above and below it, which are categorized as successive semitones on the chromatic scale. As described above, the category boundary tone is that at 50–cent deviation, and this tone is the one most difficult to assimilate into either category, that is, it is difficult to categorize the deviated tone into a note name on the chromatic scale. Therefore, it is considered that the ability to detect the tone with 50–cent deviation from a melody is the basis of the ability to encode pitches as note names.

EXPERIMENT 2–1

Method

Subjects

Subjects of this experiment were third– and fifth–grade elementary school children, first– and third–grade junior high school students, second–grade senior high school students, and university students majoring in psychology or music. All the students except the university students were divided into highly musically trained (Group H) and less well musically trained (Group L) groups. Among elementary school children, junior high school students, and senior high school students, those who had had at least four, six, and eight years, respectively, of formal training in playing the piano were placed in Group H. The remaining subjects at each education level were placed in Group L. Among the university students, the music majors were placed in Group H, and the psychology majors in Group L. The number of subjects in each group is shown below.

		Total	Group H	Group L
Elementary school				
	3rd grade (E3)	73	14	59
	5th grade (E5)	74	20	54
Junior high school				
	1st grade (J1)	73	20	53
	3rd grade (J3)	63	14	49
Senior high school				
	2nd grade (S2)	78	29	49
University				
	Psychology (U)	56		56
	Music (U)	73	73	

Materials

A total of 30 tonal melodies and 30 atonal melodies were composed (Mikumo, 1990a ; Mikumo & Umemoto, 1990 ; Umemoto, Mikumo, & Murase, 1989) ; 16, 20, and 24 melodies with lengths of 4, 5, and 6 tones, respectively, were prepared. In each melody, the pitch of one tone was deviated by 50 cents either higher or lower from that in the equal-temperament scale, and the remaining tones, which ranged from B_3 (246.94 Hz) to E_5 (659.26 Hz), were taken from the scale. The serial position where the deviated tone was placed in each melody was counterbalanced within melodies of each length. The duration of each tone was 1.0 sec (900 msec with 100-msec silence), so that the lengths of the 4-, 5-, and 6-tone melody were 4.0 sec, 5.0 sec, and 6.0 sec, respectively. The tonal melodies were in a major key, and were high in tonal melodic structure according to conventional Western rules, while the atonal melodies were low in tonal melodic structure, and both types of melodies involved a wide variety of contours.

The tones were generated by an NEC PC-9801 26K sound synthesizer board installed in an NEC PC-9801 VX personal computer, and they were recorded on tape with a VICTOR digital audio tape recorder XDZ1100.

Procedure

Subjects were instructed that this was an experiment on sense for a deviated pitch in a melody, that the melodies consisted of four, five or six tones, and that all 60 melodies without exception contained one deviated pitch from the equal–temperament scale. They were required to identify the deviated tone in each melody and to indicate their judgments by marking the position of the deviated tone in each melody on a test sheet, in which there were 60 rows of black circles, each corresponding to one of the 60 test melodies ; the number of black circles in each row corresponded to the number of tones in the corresponding melody. Three practice trials with feedback were given to the subjects prior to the 60 trials, and the interval between trials was 10 sec. The order of presentation of all 60 trials was randomized.

The melodies were played over high–quality sound reproduction equipment (CASIO digital audio tape recorder DA–1). All tones were adjusted to be equal in loudness. and were presented via two loudspeakers (SONY SRS–200).

Results and Discussion

The number of correct responses in subjects at each education level is shown in Figure 2–1, and that for tonal and atonal melodies in Groups H and L at each education level is shown in Figure 2–2.

Correct response data were analyzed in a three–way analysis of variance [6 Grades X 2 Levels of Experience in music X 2 Melody Types], with repeated measures on the third factors. The six Grades are the education levels (E3, E5, J1, J3, S2 and U), the two levels of Experience are Groups H and L, and the two Melody Types are the tonal and atonal melodies. There were significant main effects of Grade, Experience and Melody Type [$F(5,478)=40.75, p<.001$; $F(1,478)=294.29, p<.001$; $F(1,478)=848.34, p<.001$]. There were significant interactions of Grade X Melody Type, Experience X Melody Type, and Grade X Experience X Melody Type [$F(5,478)=8.25, p<.001$; $F(1,478)=32.25, p<.001$; $F(5,478)=5.88, p<.001$].

The main effect of Grade indicates that the number of correct responses increased significantly with the education level, at all levels except that between J1 and J3 ; the significance level of differences between E3 and E5, E5 and J1, J3 and S2, and S2 and U was $p<.01$. The main effect of Experience indicates that the number of correct responses in Group H was consistently higher than that in Group L. The main effect of Melody Type indicates that the number of correct responses with tonal melodies was higher than that with atonal melodies. The interaction of Grade X Melody Type indicates that, with tonal melodies, the differences among Grade in the number of correct responses were the same as those described above, whereas, with atonal melodies, the number of correct responses increased significantly with the education level, at all levels except those between E5 and J1 and between J1 and J3 ; the significance level of differences between E3 and E5, J3 and S2, and S2 and U was $p<.01$. The interaction of Experience X Melody Type indicates that the difference between Groups H and L in the number of correct responses was larger with tonal melodies than with atonal melodies.

The results suggest that the ability to detect the deviated tone in each melody depended considerably on experience or training in music. However, the result that there was no significant interaction of Grade X Experience suggests that even those who were less well trained in music acquired ability to some extent as their age increased. Moreover, the subjects in both Groups H and L showed greater accuracy in detection of the deviated tone out of a tonal melody than out of an atonal melody, and the improvement with age was more marked with tonal melodies than with atonal melodies.

In further analysis (Mikumo & Umemoto, 1990), it was found that, the later the deviated tone appeared in the melody, the grater the accuracy with which it was detected, although university music majors, could even detect a deviated tone placed at the third or fourth serial position in the melody.

If a pitch deviates from the tone in an equal–temperament scale, and can not be categorized in a permissible pitch range labeled with a note name, the pitch will be recognized as a deviated pitch. It would be rather difficult to judge whether a pitch is deviated from the equal–temperament scale without any accompanying referential tones, because the judgment would require absolute pitch. In the detection of a deviated tone out of a melody, the subject gradually constructs an internal cognitive framework (scale schema) upon hearing the tones from the beginning of the melody. This cognitive framework plays an essential role in the perceptual interpretation of each succeeding tone (Shepard, 1982a), and the subjects can detect a deviated pitch out of a melody by referring to the framework.

In this experiment, the diatonic scale schema would be constructed hearing tonal melodies, so that the detection of a deviated tone out of a tonal melody was more accurate (e.g., Krumhansl, 1979). Musically trained subjects are much more sensitive to the tonal scale structure than are untrained subjects (Krumhansl & Shepard, 1979). One might say that trained subjects can abstract the structural properties from a tonal melody, and that one effect of training is to enhance the importance of the tonal scale system in the information processing of melodies.

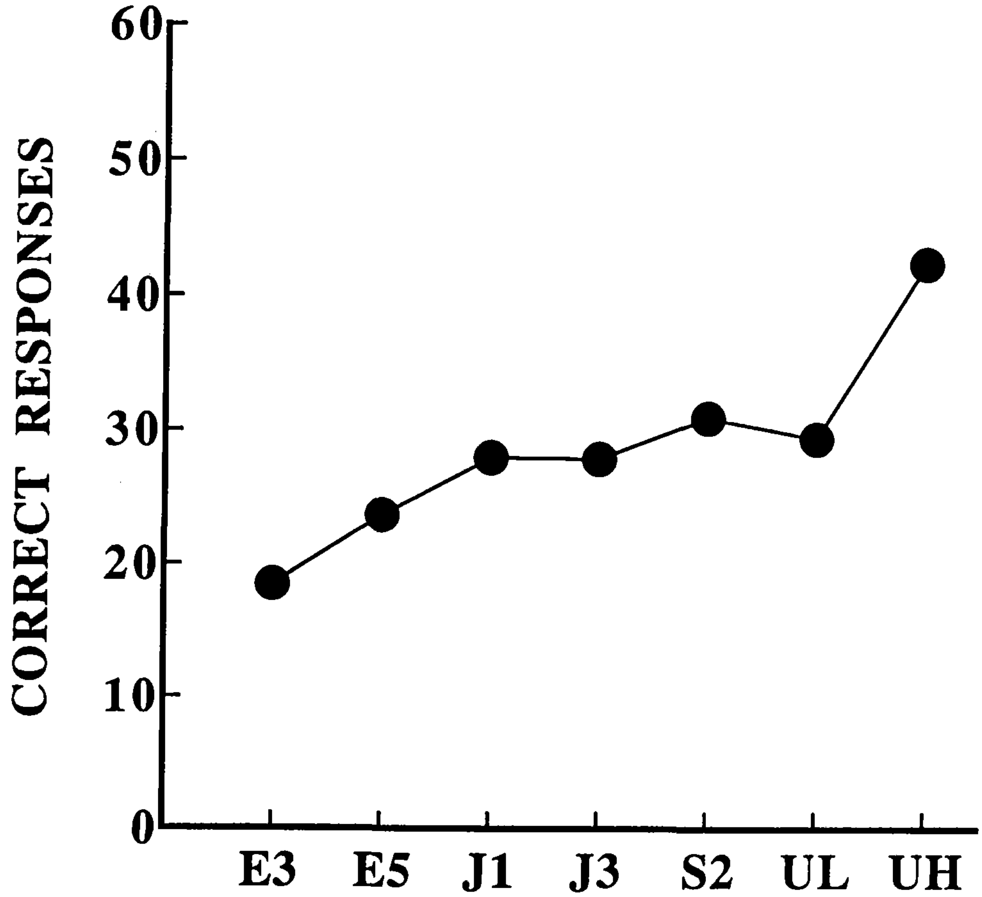


Figure 2–1.

Mean correct responses as a function of education level.

Notes.

E3 : 3rd grade of elementary school
E5 : 5th grade of elementary school
J1 : 1st grade of junior high school
J3 : 3rd grade of junior high school
S2 : 2nd grade of senior high school
UL : Less well musically trained university students
UH : Highly musically trained university students

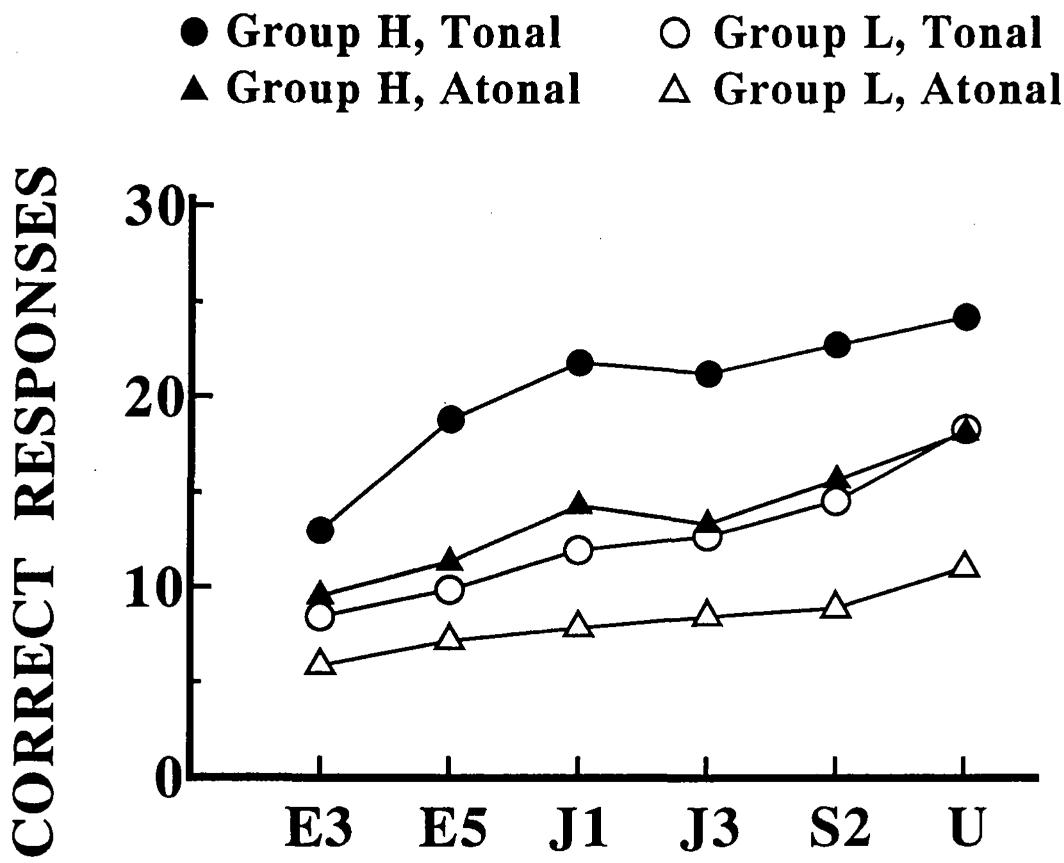


Figure 2-2.

Mean correct responses for tonal and atonal melodies in Groups H and L as a function of education level.

Notes.

E3 : 3rd grade of elementary school
E5 : 5th grade of elementary school
J1 : 1st grade of junior high school
J3 : 3rd grade of junior high school
S2 : 2nd grade of senior high school
U : University students

EXPERIMENT 2–2

The results of factorial study with 25 variables in the music tests of Wing (1948), Drake (1954), Seashore (1919) and Seashore, Lewis, & Saetveit (1960), Sherman–Knight (1977), Musical Aptitude Profile (1965), a test of absolute pitch, and the new test of detection of pitch deviation described in Experiment 2–1, are shown in Table 1. Subjects were 64 female university students, 22 of whom were music majors. The results revealed that the test of pitch deviation was one of components of the primary factor, in which other tests were also highly loaded ; these included chord analysis, chord pitches, memory for melody (Wing) ; rhythm and memory for melody (Seashore) ; pitch interval, rhythm, memory for melody, reading staff notation (Sherman–Knight) ; phrasing test (MAP) ; and absolute pitch. Therefore, the ability to detect a deviated tone from the equal–temperament scale out of a melody was confirmed to be a component of the most important factor in music cognition (Mikumo, 1990a ; Mikumo & Umemoto, 1990 ; Umemoto, 1993).

Table 1

Varimax Rotated Factor Matrix for 25 variables in the music tests of Wing, Drake, Seashore, Sherman–Knight, Musical Aptitude Profile, Absolute Pitch, and Detection of pitch deviation.

Variable	Communality	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
1 WG1	.653	**.586**	.285	.188	−.001	.297	**.325**
2 WG2	.660	**.579**	**.340**	.096	.201	**.400**	.008
3 WG3	.735	**.685**	**.483**	.074	.120	−.103	.043
4 WG4	.660	.264	**.717**	.142	.018	−.099	−.215
5 WG5	.730	.316	**.756**	.105	−.023	.178	.122
6 WG6	.642	**.348**	**.663**	.070	.073	−.234	.124
7 WG7	.760	.253	**.789**	.164	.141	.134	.099
8 DR1	.746	**−.364**	**−.698**	−.211	−.247	−.135	.048
9 DR2	.761	−.117	.100	−.065	−.125	.238	**−.813**
10 SE1	.764	.111	−.009	.059	.117	**.845**	−.147
11 SE2	.442	.075	−.003	.144	**.636**	.105	.001
12 SE3	.464	**.557**	.220	−.251	.104	.153	.089
13 SE4	.541	−.119	**.332**	.114	−.318	.116	**.521**
14 SE5	.524	−.207	.149	.177	**.660**	.082	−.041
15 SE6	.635	**.409**	**.351**	−.028	**.545**	−.169	.137
16 SK1	.398	.305	.215	**.496**	.060	−.035	−.087
17 SK2	.786	**.767**	**.321**	.241	−.156	−.074	−.086
18 SK3	.742	**.736**	.286	.059	.131	.263	.171
19 SK4	.790	**.821**	.170	.275	−.099	.038	.038
20 SK5	.687	**.566**	.270	**.505**	.019	.194	.022
21 MA1	.617	**.437**	.147	**.515**	.191	−.281	−.157
22 MA2	.605	−.011	−.039	**.715**	.165	.032	.255
23 MA3	.539	.157	.232	**.650**	.117	.146	.051
24 AP1	.811	**.852**	.146	.193	.013	−.089	−.139
25 DV1	.846	**.710**	**.383**	**.433**	.056	−.039	−.053
Contribution	.661	.234	.158	.094	.062	.059	.054

Wing Test
WG1–chord analysis, WG2–pitch change, WG3–memory,
WG4–rhythmic accent, WG5–harmony, WG6–intensity, WG7–phrasing.
Drake Test
DR1–tonal memory, DR2–rhythm.
Seashore Test
SE1–pitch, SE2–loudness, SE3–rhythm, SE4–time, SE5–timbre, SE6–tonal memory.
Sherman–Knight Test
SK1–timbre of instruments, SK2–selective attention to a melody of designated instrument 1,
SK3–rhythm identification, SK4–selective attention to a melody of designated instrument 2,
SK5–ensemble identification.
Musical Aptitude Profile Test
Musical sensitivity : MA1–phrasing, MA2–balance, MA3–style.
Absolute Pitch Test–AP1.
Detection of Pitch Deviation Test–DV1.

EXPERIMENT 3

Cerebral Hemispheric Dominance for Various Types of Melodies

The purpose of Experiment 3 was to investigate the effects of the level of musical experience and the associated encoding or processing strategies on cerebral hemispheric dominance in the processing of stimulus melodies.

There have been many studies designed to investigate the relation of each cerebral hemisphere to the processing of musical materials, but the results were not consistent with each other, as described below : Musical stimuli are processed in the right hemisphere (Damasio, Almeida, & Damasio, 1975 ; Kimura, 1964, 1967, 1968, 1973 ; Mckee, Humphrey, & McAdam, 1973 ; Milner, 1962 ; Shankweiler, 1966 ; Spellacy, 1970). During the detection of some tones from a familiar melody, the right hemisphere is dominant (Gates & Bradshaw, 1977a,b). There is no cerebral hemispheric dominance for Western melodies among highly musically trained subjects (Gordon, 1970). Singing ability is related to the right hemisphere (Bogen & Gordon, 1971). Recognition of musical chords is also related to the right hemisphere (Gordon, 1978 ; see Gordon, 1980, 1983). If a serial or sequential processing strategy is employed, the stimuli are processed in the left hemisphere, but if a parallel or simultaneous processing strategy is employed, the stimuli are processed in the right hemisphere (Cohen, 1973 ; Ohgishi, 1978). In the analytic processing strategy, stimuli are processed in the left hemisphere, while in synthetic processing strategy, they are processed in the right hemisphere (Locke & Kellar, 1973). A verbal processing system, which is specialized for speech and abstract information, may be primarily a left hemisphere function, while a visual processing system, which is more adept with spatial and concrete information, might be predominantly associated with the right hemisphere (Bower, 1970). In subjects well trained in music, a quite familiar musical stimulus is processed by the analytical processing strategy in left hemisphere, whereas, in subjects less well trained in music, musical stimulus is processed by the global processing strategy in the right hemisphere (Bever & Chiarello, 1974).

In Experiment 3, all subjects underwent a Numerals Test prior to each experiment. In this test, pairs of numerals were presented to a subject dichotically, employing the Broadbent technique (Broadbent, 1954) with Kimura's modifications (Kimura, 1961, 1963). Each trial consisted of three pairs of aurally presented numerals from 1 to 9 delivered in a male voice, with the two different numerals in the pair arriving at the two ears simultaneously (Figure 3–1). After each trial (total of six numerals), the subject was asked to report all the numerals he had heard, in any order he liked. A total of 28 trials (168 numerals) were presented to each subject, so that maximum score for each ear was 84. If the score for the right ear was significantly higher than that for the left ear, the left hemisphere was considered to be dominant for the verbal stimuli.

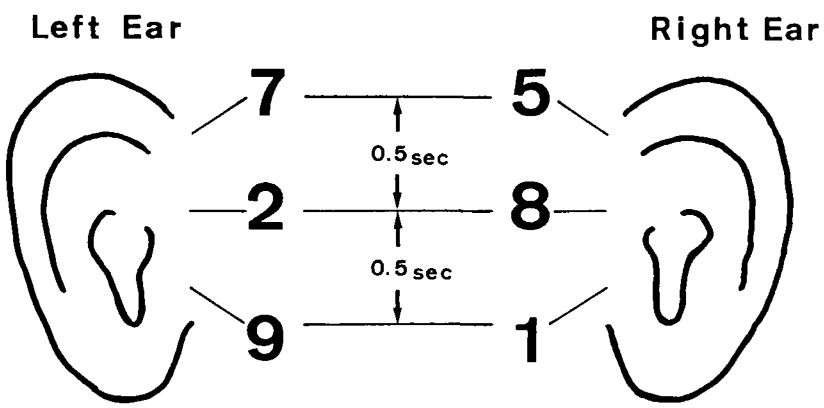


Figure 3–1.

Paradigm for dichotic presentation of numerals : Two numerals in each of three pairs are presented simultaneously for 1.0 sec to each ear, and three pairs are presented in succession separated by 0.5–sec silence.

EXPERIMENT 3–1

Cerebral hemispheric dominance for Western melodies

Method

Subjects

The highly musically trained group (Group H) consisted of 14 female subjects (average age, 21.0 years ; age range, 19–22 years) who were undergraduates majoring in music. Each had had at least 12 years of formal training in playing the piano, with an average of 13.6 years (range, 12–17 years).

The less well musically trained group (Group L) consisted of 14 female subjects (average age, 20.4 years ; age range, 19–22 years) who were undergraduates majoring in education or psychology. None were currently playing a musical instrument, and their musical experience on an instrument averaged 1.1 years (range, 0–2 years).

All subjects in Groups H and L were determined to be right–handed using the 10–item questionnaire developed by Hatta & Nakatsuka (1975), and none had hearing difficulties of which they were aware. Moreover, in the results of the Numerals Test, the score for the right ear in each subject was higher than that for the left ear, and, in both groups it was found that the left hemisphere was the dominant hemisphere for verbal stimuli.

Materials and Procedure

Eighty melodies were chosen from etude books for violin, and the tones of the melodies were generated by an NEC PC–8801 MK2 computer (Mikumo, 1987b, 1991b). Each melody was major or minor, and was a four-bar motif with the tempo arranged to be 6.0 sec in duration. After a warning signal of two binaural tones, two of the melodies matched for key, rhythm and pitch range were presented simultaneously, one to each ear, using a SONY dual-channel tape recorder and SENNHEISER stereo headphones (HD414x). After presentation of the dichotic melodies, these two melodies and two additional melodies were presented binaurally, in succession with intervals of 3.0 sec. These four binaural melodies were also matched for key, rhythm and pitch range, from G_3 (196.0Hz) to G_5 (784.0Hz). Kimura (1964) and Gordon (1970) technique was employed in Experiment 3 (Figure 3-2).

The subjects were instructed to select the two dichotic melodies from among the four binaurally presented melodies and to check the two corresponding boxes on an answer sheet during the 15.0-sec interval between trials. The positions of the two original melodies among the four choices were counterbalanced between trials and between subjects. Two practice trials with feedback were given to the subjects prior to the total of 18 test trials. Maximum score for each ear was 18. Each channel of the tape recorder was set at a comfortable standard volume. In half of the subjects in each group, the headphone channels were reversed so that any remaining asymmetries in the tape or apparatus were counterbalanced.

Results and Discussion

Recognition probability data were analyzed in a two-way analysis of variance [2 Groups X 2 Ears], with repeated measures on the second factor. There was a significant main effect of Group [$F(1,26)=5.61, p<.05$], but there was no significant main effect of Ear, and no significant interaction of Group X Ear. The results are shown in Figure 3-3.

The main effect of Group indicates that the performance of Group H was superior to that of Group L, but in both groups there was no difference between the scores for the two ears. These results suggest that there was not cerebral hemispheric dominance for Western melodies in either group, and are consistent with the findings of Gordon (1970).

The purpose of this experiment was to investigate whether musical experience had an effect on cerebral hemispheric function for processing melodies. There were two points for improvement in this experiment. One of them was the degree of familiarity with the stimuli, specifically, to what extent the subjects in each group were familiar with melodies such as those used in this experiment. Even the subjects in Group L must have often come across these sorts of tonal melodies. To investigate the effect of musical experience, it would be necessary to employ stimuli which are quite familiar to one group and quite unfamiliar to the other. This point was investigated in Experiment 3–2.

The other point for improvement was the quality of the melodies. In this experiment, these Western melodies included two musical factors: pitch and rhythm. In a lobectomy study (Milner, 1962) with the Seashore Test of Musical Ability (1919, 1960), Milner found that the right hemisphere was considerably concerned with Melody and Timbre subtests and somewhat concerned with Pitch and Loudness subtests, but was little concerned with Time and Rhythm subtests. From this point of view, it would be considered that the non–superiority of either ear observed in this experiment is due to the rhythmic aspect rather than pitch aspect. When the subjects selected the correct melodies from among the four choices, they would notice the rhythmic patterns unique to each of the dichotic pairs rather than the pitch differences. To investigate the effect of musical experience on cerebral hemispheric dominance for melodies, it would be necessary to control the factors of melodies. This point was investigated in Experiments 3–3 and 3–4.

Figure 3–2.

Paradigm for dichotic presentation of melodies : Two melodies are presented simultaneously for 6.0 sec to each ear (dichotically) followed by four successive melodies presented to both ears (binaurally), each separated by 3.0–sec silence.

EXPERIMENT 3–2

Cerebral hemispheric dominance for Japanese melodies

Method

Subjects

The highly trained group in Japanese music (Group J) consisted of six Japanese female subjects (average age, 21.4 years ; age range, 20–23 years). Each had had at least 7 years of formal training in playing the *koto* with an average of 8.7 years (range, 7–11 years).

The less well trained group in Japanese music (Group F) consisted of six foreign female subjects (average age,22.3 years ; age range, 21–24 years). They have been in Japan for 4 years at most (average years, 2.8 ; range, 1–4 years), and none were interested in Japanese music.

All subjects in Groups J and F were right–handed, and none had hearing difficulties. Moreover, in the results of the Numerals Test, the score for the right ear in each subject was higher than that for the left ear, and, in both groups it was found that the left hemisphere was the dominant hemisphere for verbal stimuli.

Materials and Procedure

Eighty melodies were chosen from Japanese folk song books, and the tones of melodies were generated by a personal computer (Mikumo, 1987b, 1991b). Each melody consisted of a Japanese scale, and was a four–bar motif with the tempo arranged to be 6.0 sec in duration. Other aspects of the method were the same as those described in Experiment 3–1.

Results and Discussion

Recognition probability data were analyzed in a two–way analysis of variance [2 Groups X 2 Ears], with repeated measures on the second factor. There was a significant main effect of Group [$F(1,10)=5.95, p<.05$], and there was significant interaction of Group X Ear [$F(1,10)=32.98, p<.001$]. The results are shown in Figure 3–4. The main effect of Group indicates that the performance of Group J was superior to that of Group F, and the interaction of Group X Ear indicates that the score for the right ear was significantly higher than that for the left ear in Group J ($p<.01$ by Newman–Keuls method), whereas the score for the left ear was significantly higher than that for the right ear in Group F ($p<.05$).

These results suggest that, for Group J, the left hemisphere was dominant for Japanese melodies, whereas for Group F, the right hemisphere was dominant for Japanese melodies. That is, the melodies familiar to the subjects in Group J were processed in their left hemisphere, whereas, the melodies unfamiliar to the subjects in Group F were processed in their right hemisphere, indicating that the subjects in each group employed different types of encoding or processing strategies for Japanese melodies. The subjects in Group J might employ the analytical processing strategy for the familiar melodies, and the subjects in Group F might employ the global processing strategy for the unfamiliar melodies. However, from these results, information concerning encoding or processing strategies cannot be obtained. In Experiments 3–3 and 3–4, the relationship between cerebral hemispheric dominance and encoding strategy for melodies was investigated.

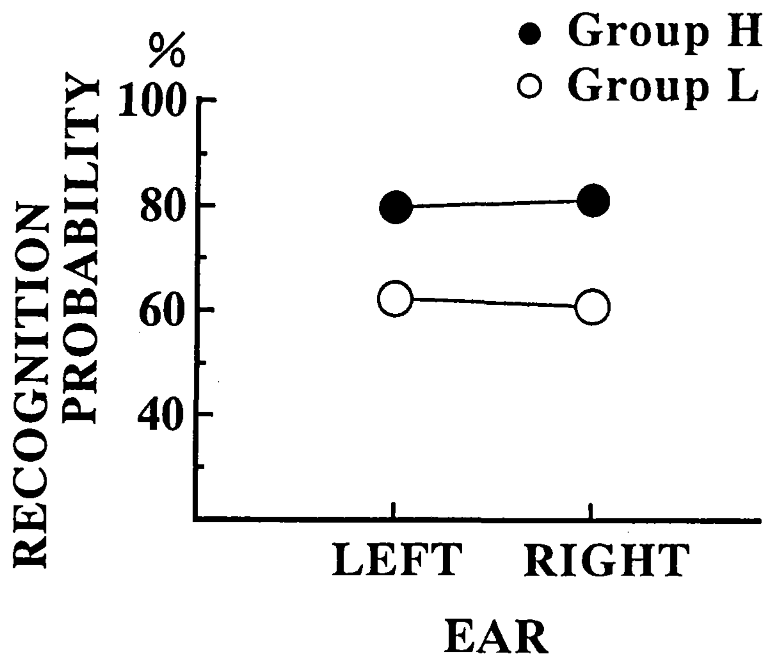


Figure 3–3.

Mean probability of correct recognition for the right and left ears in Groups H and L, hearing Western melodies.

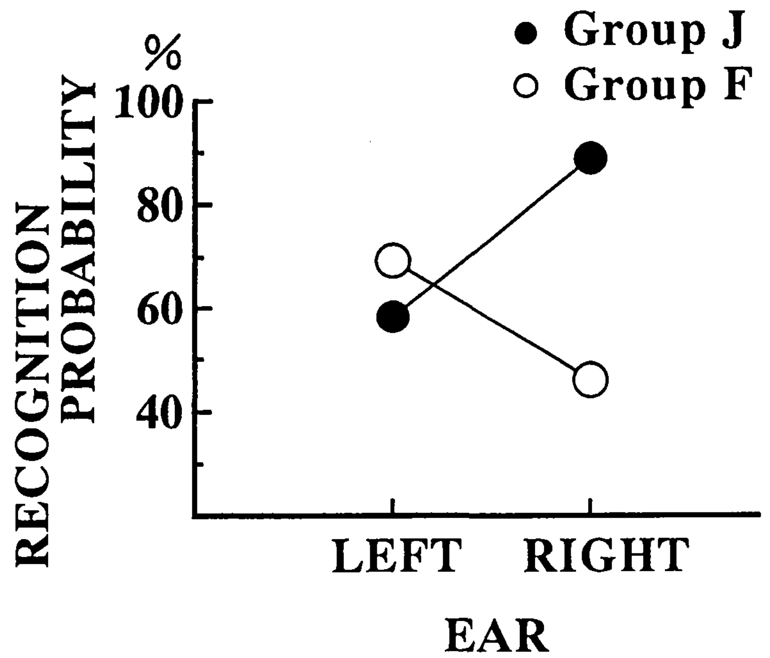


Figure 3–4.

Mean probability of correct recognition for the right and left ears in Groups J and F, hearing Japanese melodies.

EXPERIMENT 3-3

Cerebral hemispheric dominance for tonal and atonal melodies

Method

Subjects

Thirteen female subjects (average age, 21.8 years ; age range, 20–23 years) who were undergraduates majoring in music constituted the highly musically trained group (Group H). Each had had at least 13 years of formal training in playing the piano with an average of 15.8 years (range, 13–19 years).

Thirteen female subjects (average age, 21.7 years ; age range, 20–23 years) who were undergraduates majoring in domestic science constituted the less well musically trained group (Group L). None were currently playing a musical instrument, and their musical experience on an instrument averaged 2.4 years (range, 0–4 years).

All subjects in Groups H and L were right–handed, and none had hearing difficulties of which they were aware. Moreover, in the results of the Numerals Test, the score for the right ear in each subject was higher than that for the left ear, and, in both groups, it was found that the left hemisphere was the hemisphere dominant for verbal stimuli.

Materials and Procedure

Fifty–six tonal melodies and fifth–six atonal melodies were prepared (Mikumo, 1993b). Both the tonal and atonal melodies consisted of eight tones taken from an equal–temperament scale, which ranged from G_3(196.0 HZ) to E_5 (659.26 HZ). The tonal melodies were in a major key, and were high in tonal melodic structure, according to conventional Western rules, whereas the atonal melodies were low in tonal melodic structure. Both types of melodies involved a wide variety of contours. The duration of each tone was 700 msec (600 msec with 100–msec silence), so that the length of the 8–tone melody was 5.6 sec. All tones were generated by an NEC PC–9801 26K sound synthesizer board installed in an NEC PC–9801 DA computer.

After a warning signal of two binaural tones, two of the melodies were presented simultaneously, one to each ear, by a DENON digital audio dual–channel tape recorder (DTR–80P) and DENON stereo headphones (AH–D410). After the dichotic melodies, these two melodies and two additional melodies were presented binaurally, in succession with intervals of 3.0 sec. The dichotic and binaural pairs were arranged using a CANOPUS Sound Master/V with sampling frequency of 16 KHz. Two practice trials with feedback were given to the subjects prior to a total of 12 test trials, so that the maximum score for each ear was 12. Other aspects of method were the same as those described in Experiment 3–1.

Results and Discussion

Recognition probability data for Groups H and L were analyzed in a two–way analysis of variance [2 Melody Types X 2 Ears], with repeated measures on both factors. In Group H (Figure 3–5(a)), there was a significant main effect of Melody Type [$F(1,12)=28.69$, $p<.001$], and there was a significant interaction of Melody Type X Ear [$F(1,12)=37.62$, $p<.001$]. The main effect of Melody Type indicates that the performance was better with tonal melodies than with atonal melodies. The interaction of Melody Type X Ear indicates that, in tonal melodies, the score for the right ear was better than that for the left ear ($p<.01$), whereas in atonal melodies, the score for the left ear was better than that for the right ear (p<.01).

The results suggest that the tonal melodies were processed significantly better in the left hemisphere than in the right, whereas, the atonal melodies were processed significantly better in the right hemisphere than in the left. It was well recognized that the left hemisphere is specialized with respect to verbal abilities, and the right hemisphere is adapted to non-verbal tasks (cf. Brown, 1979 ; Corballis & Beale, 1976 ; Gardner, 1964 ; Gazzaniga & LeDoux, 1978 ; Geschwind & Levitsky, 1968 ; Lebrun & Zangwill, 1981 ; Norman, 1986 ; Scheid & Eccles, 1973 ; Sperry, 1982 ; Springer & Deutsch, 1985). Moreover, as found in Experiment 1, for Group H, with the high tonal structure, the pitches were readily encoded as their note names, whereas with the low tonal structure, they were not as readily encoded. Therefore, the results for Group H in Experiment 3–3 are interpreted that the tonal melodies, which are actually non–verbal stimuli in themselves but are "note name–evoking" stimuli, would be processed as verbal stimuli in the left hemisphere, and the atonal melodies would be processed as non–verbal stimuli in the right hemisphere. Thus, when the subjects in Group H memorize the tonal melodies, they would employ the verbal encoding strategy with note names.

In Group L (Figure 3–5(b)), there were significant main effects of Melody Type and Ear [$F(1,12)=13.85, p<.01$; $F(1,12)=6.37, p<.05$]. There was no significant interaction of Melody Type X Ear. The main effect of Melody Type are consistent with the findings as observed in Group H. The main effect of Ear indicates that, in both melodies, the score for the left ear was better than that for the right ear. The results suggest that the tonal and atonal melodies were processed better in the right hemisphere than in the left. Therefore, the results for Group L in this experiment suggest that the tonal and atonal melodies would be processed as non–verbal stimuli in the right hemisphere.

EXPERIMENT 3–4

Cerebral hemispheric dominance for tonal and atonal melodies along with note names

Method

Subjects

The same subjects as those in Experiment 3–3 were employed.

Materials and Procedure

Fifty–six tonal melodies and fifth–six atonal melodies were prepared (Mikumo, 1993b). Both the tonal and atonal melodies consisted of eight tones along with note names. These melodies were constructed as follows : Twelve pitches sung with note names at accurate pitches in a professional female voice were first recorded (A_3=220.00, B_3=246.94, C_4=261.63, D_4=293.66, E_4=329.63, F_4=349.23, G_4=392.00, A_4=440.00, B_4=493.88, C_5=523.25, D_5=587.33, E_5=659.26 Hz). Each of them was then cut to 600 msec, and eight pitches were chosen and were arranged serially, separated by 100 msec, for either tonal or atonal structure. These melodies, and then the dichotic and binaural pairs, were arranged using a CANOPUS Sound Master/V with sampling frequency of 32 KHz. Other aspects of the method were the same as those in Experiment 3–3.

Results and Discussion

Recognition probability data for Groups H and L were analyzed in a two–way analysis of variance [2 Melody Types X 2 Ears], with repeated measures on both factors. In Group H (Figure 3–6(a)), there were significant main effects of Melody Type and Ear [$F(1,12)=13.31, p<.01$; $F(1,12)=99.36, p<.001$], and there was no significant interaction of Melody Type X Ear. The main effect of Ear indicates that the score for the right ear was better than that for left ear for both melody types.

The results suggest that the tonal and even atonal melodies were processed significantly better in the left hemisphere than in the right. The stimuli in this experiment were the melodies sung with note names at accurate pitches, so that the subjects were given the verbal code at the same time and it was not necessary to encode pitches as verbal codes (note names) by themselves. Therefore, the subjects could encode pitches not only in tonal melodies but also in atonal melodies as verbal codes (note names).

In Group L (Figure 3–6(b)), there was a significant main effect of Melody Type [$F(1,12)=22.89, p<.001$]. When the melodies were given with note names, there was no difference between the scores for the two ears, i.e., there was no difference between the performance of the two hemispheres. This findings would suggest that the subjects attempted to use the verbal codes given, but they could not utilize them satisfactorily.

In Experiments 3–3 and 3–4, it was found that in the subjects in Group H, for not only tonal melodies but also atonal melodies, when there was a possibility that the pitches were encoded as note names, these melodies were processed in the left hemisphere.

These findings are accounted for well by a conclusion proposed by Seamon (1973) and Seamon and Gazzaniga (1973), that, when stimuli are encoded verbally and the codes are rehearsed, the left hemisphere is dominant for the processing, whereas, when the stimuli are encoded by the use of imagery, the right hemisphere is dominant for the processing ; that is, varying the encoding strategy can produce reversal of the dominant hemisphere for the stimuli (cf. Hatta, 1982 ; Silverberg, Bentin, Gaziel, Obler, & Albert, 1979 ; Silverberg, Gordon, Pollack, & Bentin, 1980 ; Yoshizaki, 1986). This phenomenon resembles the shifts found when visual material is coded verbally (Cohen, 1977).

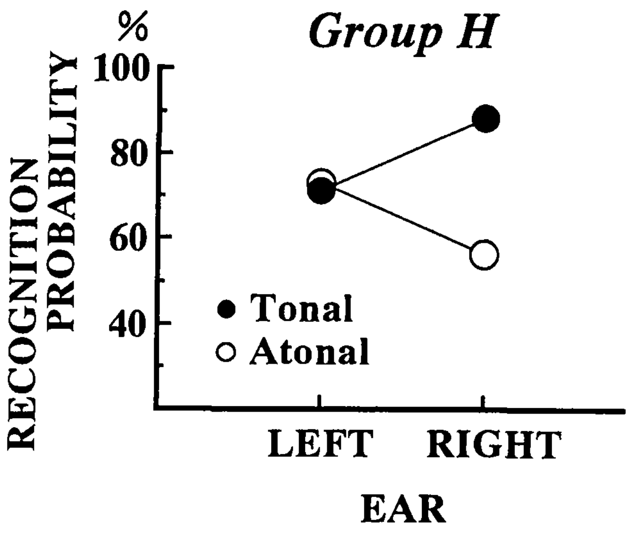


Figure 3–5 (a)
Mean probability of correct recognition for the right and left ears in Group H, hearing tonal and atonal melodies.

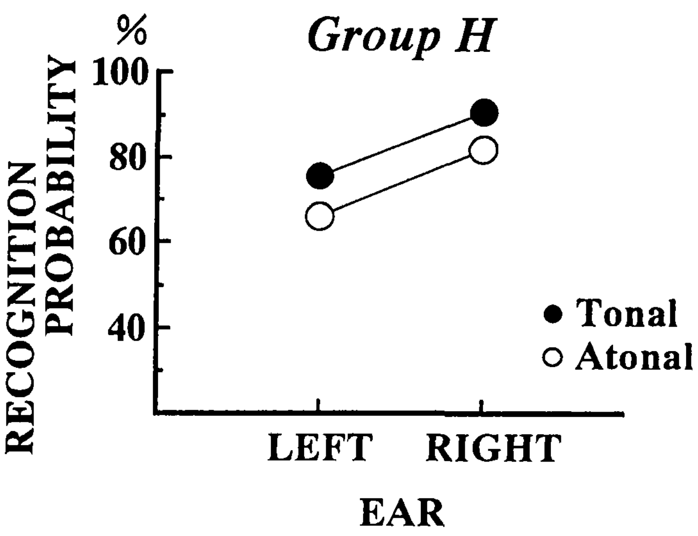


Figure 3–6 (a)
Mean probability of correct recognition for the right and left ears in Group H, hearing tonal and atonal melodies, sung with note names.

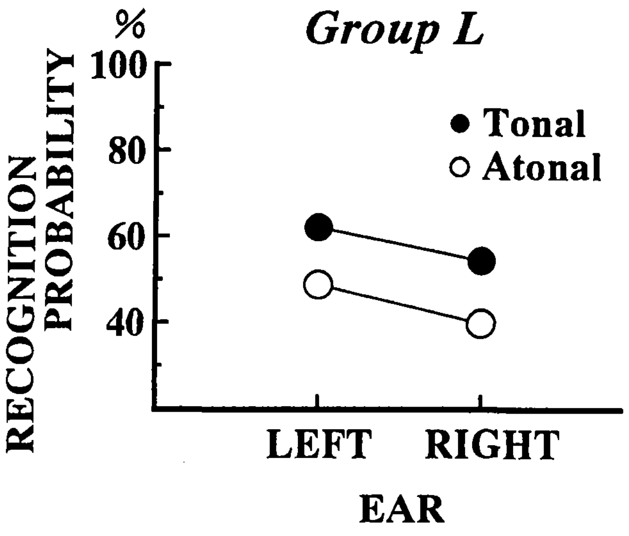


Figure 3–5 (b)
Mean probability of correct recognition for the right and left ears in Group L, hearing tonal and atonal melodies.

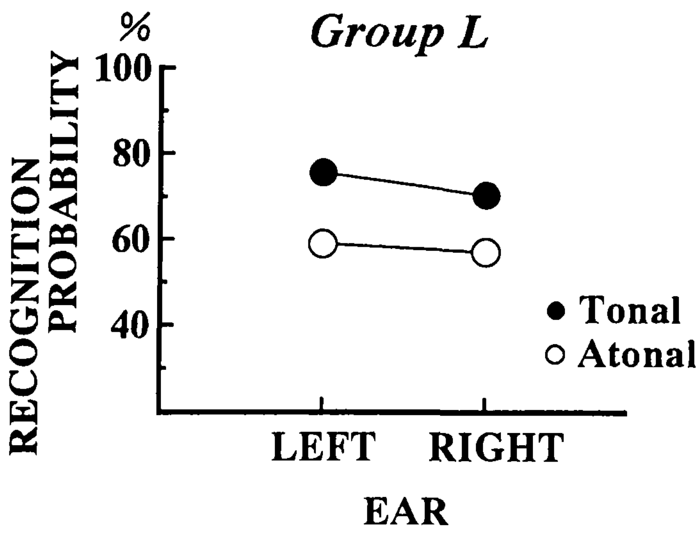


Figure 3–6 (b)
Mean probability of correct recognition for the right and left ears in Group L, hearing tonal and atonal melodies, sung with note names.

CHAPTER 3

EXPERIMENTAL TESTS OF VISUAL ENCODING STRATEGIES FOR PITCH

EXPERIMENT 4

Representations of Melodic Contour and Stuff Notation

Method

Results and Discussion

EXPERIMENT 5

Eye Tracking of Visual Representations

Method

Results and Discussion

EXPERIMENT 4

Representations of Melodic Contour and Staff Notation

In Experiment 1, it was found that melodic contour was a dominant cue for melody recognition. Many subjects, especially subjects less well trained in music, are likely to encode the pitch sequence of a short melody as a melodic contour. On the assumption that various visual and auditory contours are perceived in broadly similar ways, there is a possibility that subjects would visualize an auditory contour (auditory imagery), which may be acquired from the pitch sequence of the stimulus, as a visual contour (visual imagery). Some highly musically trained subjects reported (Chapter1) that they imaged staff notation rather than melodic contour. The purpose of Experiment 4 was to investigate whether the visualizing strategies were employed to encode pitches. This intermodal analogy may be achieved by regarding the visual abscissa as auditory time and the visual ordinate as the pitch height of the tone.

In the preliminary demonstration (Chapter 1), subjects reported that the visualization of auditory imagery as visual imagery of contour or notation did not occur simultaneously but successively ; they did not image the total configuration of the contour or notation at the same time, but from the left to the right of the configuration little by little, analogous to reading a staff notation. Subjects also reported that, to retain their images, they either rehearsed the visual image or, once the visual image was constructed (Hebb (1968) proposed that eye movements are supposed to be necessary for image construction), they tracked or scanned it (e.g., Kosslyn, 1973, 1975, 1980 ; Neisser, 1967).

Considering these subjects' reports, in Experiment 4, a visual tracking task was prepared, which is analogous to tracking the visual image, and was interpolated during the retention interval between standard and comparison melodies (Mikumo, 1993a). Subjects were required to judge whether the standard and comparison melodies were the same or different and to track the visual task. The visual tracking task included some lure patterns, therefore,it was hypothesized that the subject's involuntary matching between visual task–tracking and mental image–tracking would occur during the retention interval (see Brooks, 1968). Thus, facilitation by the correct visual patterns and disruption by the lure patterns[3] of the recognition performance compared to that in the control condition of blank–retention interval patterns would imply that the subjects had retained accurate visual representations for auditory information, and that the visualizing strategy is effective for encoding pitches of melodies.

Method

Subjects

Twelve female subjects (average age, 22.4 years ; age range, 19–24 years) who were undergraduate and graduate students majoring in music constituted the highly musically trained group (Group H). Each had had at least 15 years of formal training in playing the piano, with an average of 16.2 years (range, 15–19 years).

Twelve female subjects (average age, 21.8 years ; age range, 19–23 years) who were undergraduate and graduate students majoring in education or psychology constituted the less well musically trained group (Group L). None were currently playing a musical instrument, and their musical experience on an instrument averaged 3.4 years (range, 0–5 years).

Materials

Each trial involved a standard series followed by a retention interval and then by a comparison series. Both the standard and comparison series consisted of eight tones taken from an equal–temperament scale, which ranged from A_3 (220.0 Hz) to E_5 (659.26 Hz). The duration of each tone was 700 msec (600 msec with 100–msec silence), so that the length of the 8–tone series was 5.6 sec.

For this experiment, three lists were prepared ; one list was used for one visual condition described below, and each list involved 32 trials. Half of them (16 trials) were tonal melodies in a major key, and were high in tonal melodic structure according to conventional Western rules. The other half (16 trials) were atonal melodies, and were low in tonal melodic structure were used. Both types of melodies involved a wide variety of contours. Half of the comparison series were exactly the same as the standard series, and the other half were different from the standard.

Either of two visual tracking conditions, Staff Notation and Melodic Contour, was interpolated during the retention interval between the standard and the comparison series. The Staff Notation consisted of a series of eight musical notes on a staff without a G clef, and was presented on a computer display. The eight notes appeared successively from the left to the right with a constant space interval of about 2.5 cm on the staff (line interval : 2.0 cm). The Melodic Contour was an ascending and descending pattern of melody, and was represented as eight circles with a diameter of 0.5 cm connected by seven line segments[4], which also appeared successively from the left to the right on the computer display.

Each visual tracking condition included four patterns. "Pause" pattern (**P**) indicates a blank retention interval. "Same" pattern (**S**) indicates that the movement of the notes or circle was exactly the same as that of the 8–tone standard series ; the visual distances between the notes or circles exactly corresponded to the auditory pitch intervals. "Con–

tour-preserving" pattern (**C**) was obtained by changing one of the pitches by two semitones (higher or lower), so that the exact melodic intervals were not preserved, but with preservation of the contour of the standard series ; the visual distances between the notes or circles were somewhat different from the auditory pitch intervals. "Retrograde" pattern (**R**) was obtained by reversing the order of pitches of the standard series, so that neither the melodic intervals nor the contour were completely preserved ; the visual distances between the notes or circles were completely different from the auditory pitch intervals.

In the visual tracking task, the eight notes or eight circles appeared successively with a constant duration of 500 msec[(5)] in a series, and the series (4.0 sec) was repeated three times during the retention interval. Therefore, each trial involved a 5.6–sec standard series followed by a 2.0– sec interval, a 12.0–sec visual tracking task, a 2.0–sec interval, and then a 5.6–sec comparison series.

The tones of the standard and comparison series were generated by an NEC PC–9801 26K sound synthesizer board installed in an NEC PC–9801 DA personal computer, recorded on tape, and presented over high–quality sound reproduction equipment (DENON digital audio tape recorder DTR–80P).

Procedure

Each subject sat in front of a personal computer and wore headphones, and all tones were adjusted to be equal in loudness. In half of the subjects began, the Staff Notation condition preceded the Melodic Contour condition, and in the other half the Melodic Contour condition preceded the Staff Notation condition. The subjects were instructed that this was an experiment on memory for melodies, and that in each of the 32 trials in each visual condition, they would first hear a trial number, then a first melody (standard series), followed by a visual tracking task in which 24 notes or circles would appear corresponding to three times the 8–tone first melody, and then a second melody (comparison series). The subjects were required to judge whether the two melodies were the same or different in pitch, and to indicate their judgments by writing "S" (Same) or "D" (Different) on an answer sheet. The subjects were also instructed that they should track the movement of notes or circles presented during the retention interval but should not employ any motor encoding strategy analogous to playing the piano. The second condition was performed thirty minutes after the first condition. Three practice trials with feedback were given to the subjects prior to the 32 trials ; the interval between trials was 15 sec.

In each visual tracking condition in this experiment, there were 16 tonal trials and 16 atonal trials. In both types of melodies, four trials were run for each pattern. The four patterns (P, S, C, R) of the visual condition were alternated, and the order of presentation of all 32 trials was randomized. As described above, for this experiment, three lists were prepared, and each list involved 32 trials. These lists were counterbalanced between subjects to minimize the differences among the trials in each pattern under the two visual conditions in difficulty.

Results and Discussion

The results in Groups H and L are shown in Figure 4. Recognition probability data (hit rate plus correct rejection rate) were analyzed in a four-way analysis of variance [2 Groups X 2 Visual Conditions X 2 Melody Types X 4 Visual Patterns], with repeated measures on the second, third and fourth factors. There were significant main effects of Group, Melody Type and Visual Pattern [$F(1,22)=38.49, p<.001$; $F(1,22)=64.20, p<.001$; $F(3,66)=4.00, p<.05$]. There were significant interactions of Group X Melody Type, Group X Visual Pattern and Group X Visual Condition X Visual Pattern [$F(1,22)=4.52, p<.05$; $F(3,66)=4.75, p<.01$; $F(3,66)=2.74, p<.05$]. The main effect of Group indicates that the performance of Group H was consistently superior to that of Group L. The main effect of Melody Type indicates that the recognition probability with tonal melodies was higher than that with atonal melodies. The main effect of Visual Pattern indicates that there were significant differences among the recognition probabilities in four visual patterns.

To obtain detailed results concerning the interactions in the four-way analysis presented above, recognition probability data for Groups H and L were analyzed in a three-way analysis of variance [2 Visual Conditions X 2 Melody Types X 4 Visual Patterns], with repeated measures on all three factors. For Group H, there was a significant main effect of Melody Type [$F(1,11)=70.52, p<.001$], and there was a marginal interaction of Visual Condition X Visual Pattern [$F(3,33)=2.49, p=.076$]. For Group L, there were significant main effects of Melody Type and Visual Pattern [$F(1,11)=13.53, p<.01$; $F(3,33)=6.29$, $p<.01$]. The interaction of Group X Melody Type in the four-way analysis indicates that the difference between the performance with tonal and atonal melodies in Group H ($p<.001$) was greater than that in Group L ($p<.01$). The marginal interaction of Visual Condition X Visual Pattern in Group H indicates that, when Staff Notation was interpolated during the retention interval, the performance for C was significantly lower than that for S ($p<.01$), R and P (both $p<.05$), whereas, when Melodic Contour was interpolated, there were no significant differences among the visual patterns (by Newman–Keuls method). The main effect of Visual Pattern in Group L indicates that, the performance for R was significantly lower than that for C, S, and P (all $p<.05$).

The results suggest that, for Group H in Staff Notation condition, visual tracking of S pattern somewhat facilitated the recognition performance, whereas tracking of C pattern significantly disrupted the performance compared to that with the control pattern of P ; S was compatible with the subject's visual image but C was incompatible with it. The results imply that the subjects had retained an accurate representation of staff notations ; that is, the visual distance between notes in their internal representations precisely reflected the auditory pitch intervals of the standard melodies. Visualization of auditory imagery as staff notations would be an effective strategy to encode the pitches of melodies. Although R was completely different from the standard series, the performance was as good as that in P, implying that the subjects would be diverting attention away from the R pattern and employing not visualization but another strategy. More detailed evidence for this hypothesis was obtained in the study of eye movements in Experiment 5.

For Group H in Melodic Contour condition, compared to the control pattern of P, there were no significant differences among the four visual patterns, i.e., the performance was not influenced by any visual tracking patterns in Melodic Contour. This result implies that these subjects had not retained visual representation of melodic contours as used in this experiment, and that in these subjects visualization of auditory imagery as melodic contours would be an ineffectual strategy to encode pitches.

On the other hand, for Group L in both visual conditions, the recognition performance in S and C patterns were somewhat better than, or were as good as, that in the control pattern P, whereas tracking of R pattern significantly disrupted the performance compared to that in P pattern ; S and C patterns were compatible with the subject's visual images but R pattern was completely incompatible with it. The similarity of the recognition performance in both visual tracking conditions would imply that the subjects abstract the melodic contour from the note sequence in the Staff Notation condition. These results imply that, in both visual tracking conditions, the subjects had retained visual representations of melodic contours but with decreased accuracy, since their performance was not influenced by the C visual tracking pattern ; that is, the visual distance between notes or circles in their internal representations roughly reflected the auditory pitch intervals of the standard melodies. Visualization of auditory imagery as melodic contours would be to some extent an effective strategy for encoding pitches of melodies. The findings support the view that untrained subjects do not find contour recognition much more difficult than trained subjects (e.g., Bartlett & Dowling, 1980), but find pitch interval recognition more difficult (e.g., Cuddy & Cohen, 1976 ; Bartlett & Dowling, 1980).

Davies and Jennings (1977) reported that musically unsophisticated listeners can make fairly accurate drawings of the contours of melodies they have just heard.

The findings in Experiment 4 are consistent with those in Experiment 1 : whereas the subjects in Group H were able to encode pitches as accurate notes on a staff, the subjects in Group L encoded the pitch sequence as a contour. It is quite evident that there is an inter–modal analogy between the perception of pitch relationships and that of relationships in visual space.

In this experiment, some subjects muttered the pitches or the note names to themselves in a whisper during the visual tracking task. It was hypothesized that some subjects would attempt to employ a dual encoding strategy ; for example, a pitch rehearsal or verbal encod–ing strategy along with the visual encoding strategy. This hypothesize was investigated in Experiment 6.

Visual imagery is identified as one component of a working memory system (cf. Badde–ley, 1986, 1990 ; Baddeley & Lieberman, 1980). Baddeley (1986) proposed that an eye movement system may be used to rehearse and maintain the image ; therefore the purpose of Experiment 5 was to investigate the eye movement system during a visual tracking task.

****

(3) In this case, visual task–tracking interferes with mental image–tracking. The concep–tual design of this task is similar to that of Brooks's image scanning task, in which a visual task is proposed to interfere with a visual image (Brooks, 1968).

(4) Melodic Contour was designed to consist of not only seven line segments but also eight small circles, which appeared as the corners of the melodic contour corresponding to the eight notes. The circles were included because Baker and Loeb (1973) have provided some support for the importance of corners in visual perception ; they recorded eye movements and found that their observers fixated for longer periods on the corners of the patterns than on the other parts.

(5) Various durations were presented in the preliminary experiment, and subjects responses indicated that 500 msec was the most appropriate duration for visual tracking.

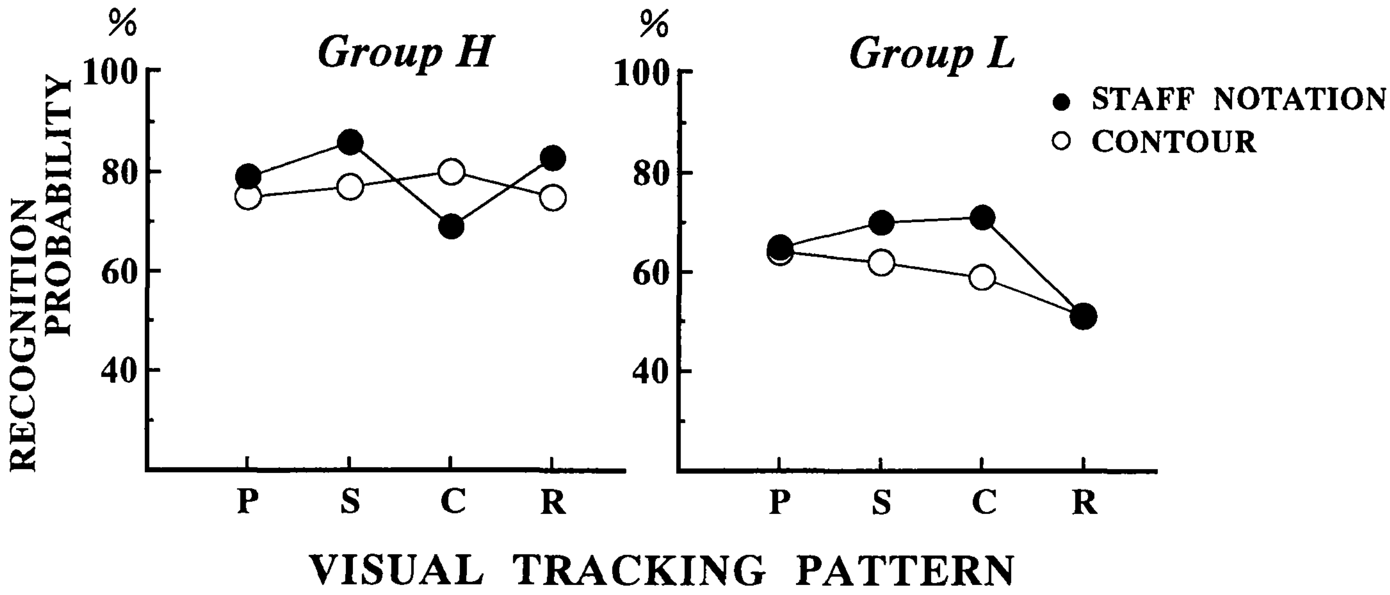


Figure 4.

Mean probability of correct recognition (hit rate plus correct rejection rate) for the four visual tracking patterns (P, S, C, R) in Staff Notation and Melodic Contour conditions in Groups H and L.

Notes.

P : Pause, S : Same (as the standard series),

C : Contour–preserving, R : Retrograde.

EXPERIMENT 5

Eye Tracking of Visual Representations

Baddeley (1986) obtained evidence that concurrent eye movements had a considerably disruptive effect on performance in the formation and utilization of a visuo–spatial image. Therefore, Baddeley proposed the possibility that eye movements reflect a process involved in rehearsing or maintaining the image, in a way analogous to that in which articulation is assumed to maintain the phonological short–term trace. His result would be indirect evidence that the eye movement system is used to rehearse and maintain the visuo–spatial image. Hebb (1968) has proposed that eye movements have an essential organizing function in visuo–spatial image.

In Experiment 5, in which the method was almost the same as that in Experiment 4, the eye movements during the retention interval between the standard and comparison melodies were recorded using an eyemark recorder. The visual tracking task interpolated during the retention interval was prepared as the external representation analogous to tracking the visual image which would be constructed by the auditory information of the standard melody. There is a reciprocal relationship between the processing of external (percepts) and internal (imagery) stimuli (Marks, 1973). Subjects were required to make a recognition judgment of melodies and to track the visual task. The visual tracking task included some lure patterns, therefore, it was hypothesized that the subject's involuntary matching between the visual task–tracking and mental image–tracking would occur during the retention interval (cf. Brooks, 1968). Thus, accurate visual tracking in response to the correct visual patterns, and deviated, confused or backward visual tracking in response to the lure visual patterns[3] would imply that the subjects had retained accurate visual representations for auditory information, and that they precisely tracked their visual image during visualizing.

Method

Subjects

Six female subjects (average age, 21.9 years ; age range, 19–23 years) who were undergraduate and graduate students majoring in music constituted the highly musically trained group (Group H). Each had had at least 15 years of formal training in playing the piano, with an average of 16.9 years (range, 15–19 years).

Six female subjects (average age, 22.1 years ; age range, 20–24 years) who were undergraduate and graduate students majoring in psychology constituted the less well musically trained group (Group L). None were currently playing a musical instrument, and their musical experience on an instrument averaged 2.5 years (range, 0–4 years).

All subjects had good eyesight without glasses or contact lenses. None of the subjects were those employed in Experiment 4.

Materials

Each trial involved a standard series followed by a retention interval with a visual tracking task and then by a comparison series. Both the standard and comparison series consisted of eight tones. For this experiment, two visual tracking conditions were prepared, and each condition involved 12 trials. Half of them (6 trials) were tonal melodies in a major key, and the other half (6 trials) were atonal melodies. Of the comparison series, half were exactly the same as the standard series, and the other half were different from the standard.

Either of two visual tracking conditions was interpolated during the retention interval. The Staff Notation condition consisted of a series of eight musical notes on a staff. The eight notes appeared successively from the left to the right on a computer display. The Melodic Contour condition was an ascending and descending pattern of melody, and consisting of eight circles with a diameter of 0.5 cm connected by seven line segments(4). The eight circles and the seven line segments also appeared successively from the left to the right on a computer display.

Each visual tracking condition included three patterns. "Same" pattern **(S)** indicated that the movement of the circle was exactly the same as that of the 8–tone standard series ; the visual distances between the notes or circles exactly corresponded to the auditory pitch intervals. "Contour–preserving" pattern **(C)** was obtained by changing two pitches, but preserving the contour of the standard series ; the visual distances between the notes or circles were somewhat different from the auditory pitch intervals. "Retrograde" pattern **(R)** was obtained by reversing the order of pitches of the standard series ; the visual distances between the notes or circles were completely different from the auditory pitch intervals. The order of presentation of all 12 trials in each condition was randomized.

In the visual tracking task, the eight notes or circles appeared successively with a constant duration of 500 msec(5) in a series, and the series (4.0 sec) was repeated twice during the retention interval. In the series, the subjects first saw one note or circle, followed by notes or circles which were added one by one until they saw eight notes or circles ; a brief blank interval was then interpolated, and the series was then repeated. Each trial involved a 5.6–sec standard series followed by a 2.0–sec interval, an 8.0–sec visual tracking task, a 2.0–sec interval, and then a 5.6–sec comparison series.

Procedure

Each subject sat in front of a personal computer and wore the head unit[6] of an eye mark recorder (NAC Eye Mark Recorder model EMR-600[7]) ; the distance between the computer display and the elliptical mirror unit was about 80 cm. All tones were adjusted to be equal in loudness (approximately 50 dB SPL), and were presented via a loudspeaker. Subjects were instructed that this was an experiment on memory for melodies, and that in each of the 12 trials in each visual tracking condition, they would first see a trial number, then hear a first melody (standard melody), followed by a visual tracking task in which 16 notes or circles would appear corresponding to twice the 8-tone first melody, and then a second melody (comparison series). The interval between trials was 15 sec. The subject's task was to judge whether the two melodies were the same or different in pitch and to answer orally. The subjects were also instructed that they should track the movement of notes or circles, but should not employ any motor encoding strategy analogous to playing the piano. After fine adjustments were made to the head unit, the eyeball movement characteristics for the individual subject were calibrated[8]. Three practice trials with feedback were given to the subjects prior to the 12 trials in each condition. The other aspects of the method were the same as those in Experiment 4.

Results and Discussion

In this experiment, the subjects were to judge whether the standard series and the comparison series were the same or different in pitch ; however, the purpose of this experiment was not to analyze the recognition probability, but to investigate the eye movements during the retention interval.

In the three patterns of two melody types in each condition, the eye–fixation duration on each of 16 notes or 16 circles in the pattern corresponding to twice the 8–tone standard series were measured. In each visual pattern, the fixation durations on eight notes or circles in the latter half were available, because at the early stage of the retention interval, the subjects' eye movements were not stable. In Figures 5–1 and 5–2, the average duration of fixation as a function of the serial position in the latter half of each pattern is shown in each group. The data were obtained from the subject's dominant eye. The samples of eye tracking data in the latter half of each pattern for the subjects in Groups H and L are also shown in Figures 5–3 and 5–4. Fixation durations of more than 100 msec with visual angles of less than 2.0 degree are marked by circles, and sequential fixations are connected by straight lines ; no attempt has been made to preserve the original saccadic path. The numbers shown in parentheses in the text below correspond to those in Figures 5–1 to 5–4.

Either of the visual tracking conditions (Staff Notation or Melodic Contour) was interpolated during the retention interval, while the subjects attempted to encode the pitches of the standard series. When the 8 notes or 8 circles appeared at compatible points which were consistent with the subject's internal visual representations, it was hypothesized that it would be easier to track and fixate them, therefore, the fixation durations on them would be longer, and there would be not marked differences among them. The notes or circles appeared with a constant duration of 500 msec at each point, so that the fixation duration on each point would be 500 msec at most. On the other hand, when some notes or circles appeared at incompatible points which were inconsistent with the subject's internal visual representations, it was hypothesized that either it would be difficult to track and fixate them for a long duration, or, if the subjects showed a response to the novelty of the points, then the fixation durations would be rather longer on the unexpected points than on the others or backward tracking to the unexpected points would occur.

For the subjects in Group H, in the Staff Notation condition (Figure 5–1), as a control trial, eight notes were presented successively with a constant duration of 500 msec on a staff without the standard and comparison melodies. The fixation duration on each of them was about 250 msec.

For tonal melodies, in S (No.1, 2), in which all eight notes appeared at the exact positions on a staff as the pitches of the standard series, there were not marked differences among the fixation durations on the eight notes. The duration on each of them was about 420 msec and was considerably longer than in the control trial, indicating that it was easy to track and fixate all eight notes which appeared at expected positions.

In C, in which the fourth and sixth notes (No.3) or the second and fifth notes (No.4) appeared at incorrect positions on a staff, the fixation durations were longer on these two notes than on the others, indicating that the subjects were responding to the novelty of the unexpected positions at which these two notes appeared. The eye tracking data (Figure 5–3) revealed back–tracks[9] to the unexpected notes. The fixation durations on these two notes included the re–fixation (back–track) duration. The results support the finding obtained in Experiment 4 that recognition performance was disrupted by C pattern.

In R (No.5, 6), in which all eight notes appeared at incorrect positions on a staff, there were not marked differences among the fixation durations on the eight notes. The duration on each of them was about 160 msec (No.5) or 200 msec (No.6) and was shorter than in the control trial, indicating that it was difficult to fixate all eight notes which appeared at unexpected positions. The eye tracking data (Figure 5–3) revealed more unstable and confused tracking (zigzag–track) in R than in the control trial. However, the recognition performance in Experiment 4 was not markedly disrupted by R pattern, implying that the staff notation which consisted of incorrect eight notes was so different from the standard series that the subjects showed little concern about the precise positions where all eight notes appeared. That is, since visualizing staff notation was interfered with R pattern, they might switch to another encoding strategy in a different modality.

For atonal melodies, in S (No.7, 8), there were not marked differences among the fixation durations on all eight notes. The duration on each of them was about 370 msec. In C, in which the third and seventh notes (No.9) or the fifth and seventh notes (No.10) appeared at incorrect positions on a staff, the fixation durations were longer on these two notes than on the others. The eye tracking data (Figure 5–3) also revealed back–tracks to the unexpected notes with atonal melodies. In R, there were not marked differences among the fixation durations on all eight note (No.11) or on the third to eighth notes (No.12). The duration on each of them was about 170 msec and was shorter than in the control trial. In S, C and R, the fixation duration was consistently shorter than that with tonal melodies, but the pattern was similar to that with tonal melodies.

These findings indicate that the subjects in Group H could visualize pitches as accurate notes on a staff to encode pitches, and that they precisely tracked the internal visual representations of notes. The visual distances between notes in their internal representations precisely reflect the auditory pitch intervals of the standard melodies.

On the other hand, in the Melodic Contour condition (Figure 5–2), as a control trial, eight circles connected by line segments were presented successively with a constant duration of 500 msec, without the standard and comparison melodies. The fixation duration on each of them was about 290 msec and was somewhat longer than that in the Staff Notation condition.

For tonal and atonal melodies, in S (No.1, 2, 7, 8), the visual distances between circles were exactly the same as the auditory pitch intervals. In C, the fifth and sixth circles (No.3), the third and seventh circles (No.4), the third and sixth circles (No.9), or the second and sixth circles (No.10) appeared at somewhat deviated points from the exact points, and the visual distances between circles were somewhat different from the auditory pitch intervals.

In R (No. 5, 6, 11, 12), the visual distances between circles were completely different from the auditory pitch intervals (No. 5, 12) or were somewhat similar to the auditory pitch intervals (No. 6, 11).

In S, C and even R with tonal and atonal melodies, there were not marked differences among the fixation durations on all eight circles in the pattern, and the duration on each of them was about 300 msec, although in R with atonal melodies it was somewhat shorter than in the other patterns. This would indicate that, whichever of these patterns appeared on the computer display, the subjects could track well (Figure 5–4) and fixate the circles to the same extent as in the control trial ; the subjects depended not on visualization of pitches as the melodic contour but on another strategy in a different modality, although in atonal melodies they somewhat depended on representations of melodic contours to encode pitches. These results support the finding obtained in Experiment 4 that recognition performance was not facilitated or disrupted by any pattern of visual tracking task in the Melodic Contour condition.

For the subjects in Group L, in the Staff Notation condition (Figure 5–1), as a control trial, eight notes were presented successively with a constant duration of 500 msec on a staff, without the standard and comparison melodies. The fixation duration on each of them was about 220 msec and was somewhat shorter than that in Group H.

For tonal and atonal melodies, in S (No.1, 2, 7, 8), there were not marked differences among the fixation durations on the eight notes. The duration on each of them was about 300 msec for tonal melodies (No. 1, 2) and about 270 msec for atonal melodies (No.7, 8), and somewhat longer than that in the control trial. In C, the fourth and sixth notes (No.3), the second and fifth notes (No.4), the third and seventh notes (No.9), or the fifth and seventh notes (No.10) appeared at incorrect positions on a staff, but there were not marked differences among the fixation durations on all eight notes, and they were almost the same as those in S pattern. In R, the fixation durations on the second to eighth notes were considerably shorter, when the melodic contour formed by the note sequence in the visual tracking task was completely different from that of the standard melody (No. 5, 11) ; the duration on each of them was less than 200 msec. However, when the melodic contour formed by the note sequence in the visual tracking task was similar to that of the standard melody (No. 6, 12), the fixation durations were relatively longer.

On the other hand, in the Melodic Contour condition (Figure 5–2), as a control trial, eight circles were presented successively with a constant duration of 500 msec connecting line segments without the standard and comparison melodies. The fixation duration on each of them was about 290 msec and was almost the same as that in Group H.

For tonal and atonal melodies, in S (No.1, 2, 7, 8), there were not marked differences among the fixation durations on all eight circles. The duration on each of them was about 320 msec. In C, the fifth and sixth circles (No.3), the third and seventh circles (No.4), the third and sixth circles (No.9), or the second and sixth circles (No.10) appeared at somewhat deviated points from the exact points, but there were not marked differences among the fixation durations on the eight circles, and they were almost the same as those in S pattern. In R (No.5, 6, 11, 12), the fixation durations on the second to eighth circles (No.5, 6, 12) or on the fourth to eighth circles (No.11) were considerably shorter. The duration on each of them was less than 200 msec.

In Group L, for either visual tracking condition with either melody type, there were not marked differences among the fixation durations on all eight notes or circles in S, C or R. The similarity of the fixation data in both visual tracking conditions would imply that the subjects abstract the melodic contour from the note sequence in the Staff Notation condi–

tion. The fixation durations on each note or circle in S and C were almost the same, indicating that in C pattern, in which the notes or circles appeared at somewhat deviated positions from the standard melody, the subjects showed little concern about the small positional deviations. On the other hand, the fixation durations on each note or circle were considerably shorter in R than in S and C, indicating that the subjects' eye movements were considerably disrupted by R pattern, in which the notes or circles appeared at completely different positions from the standard melody. In both visual tracking conditions, the eye tracking data in S and C patterns revealed smooth tracking movements, but that in R pattern revealed zigzag tracks (Figures 5–3 and 5–4). Therefore, it was found that the subjects in Group L depended considerably on the visualization of pitches as the melodic contour, and that they roughly tracked their internal visual representations of the melodic contour. The visual distances between notes or circles in their internal representations roughly reflect the auditory pitch intervals of the standard melodies. These results support the finding obtained in Experiment 4 that recognition performance was disrupted by R patterns in either visual tracking condition.

Goolsby (1989) reviews the procedures and equipment used in previous eye movement studies in music reading. When reading music, eye movements determine *what* portion of the music notation is made available for cognitive processing by the reader, eye movements determine the *order* in which this music notation is made available to the reader, and eye movements determine the *duration* that this information is made available to the reader (McConkie, 1983).

****

Footnotes (3) (4) and (5) —— see Experiment 4.

(6) The head unit, which is mounted on the subject's head, consists of an Eye Mark Sensor unit for the detection of eye marks of both eyes, a Field Shooting Camera Head for shooting of field view, an Elliptical Mirror Unit for collection and focusing of the light-emitting diode (LED) light image reflected from the eyeball, and a Goggle Unit which holds this head units.

(7) The system employs both an X–Y coordinate separated optical system with a rotational ellipsoidal mirror and a special scanning MOS sensor for unique eye mark detection. It can sample the eye mark data of both eyes at the resolution of 0.17–degree minimum angle readout at the sampling frequency of 600 Hz. It has a high–quality sensor to detect subject's head motion.

(8) Calibration is required to eliminate the measurement distortion due to differences between subjects and optical distortion.

(9) More direct evidence has been obtained in studies of eye movements during reading. Carpenter and Just (1977) noted that, in reading ambiguous sentences, the subjects' eyes often back–tracked to precisely the source of ambiguity. The result was interpreted as indicating that the subjects had retained an accurate spatial representation of at least the last few words.

Staff Notation

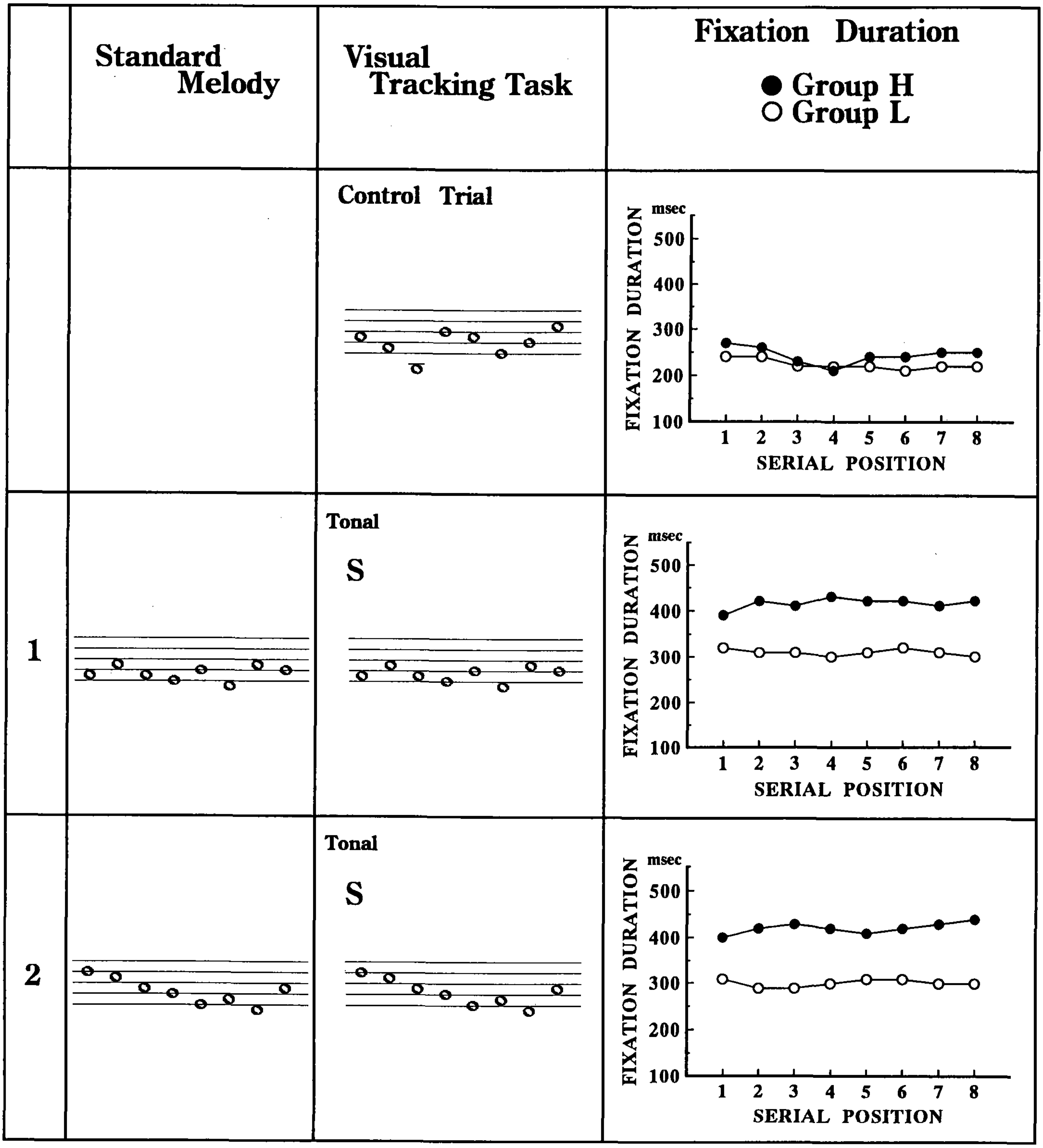
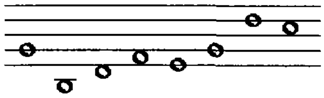
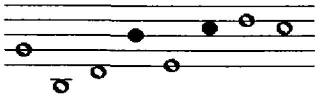
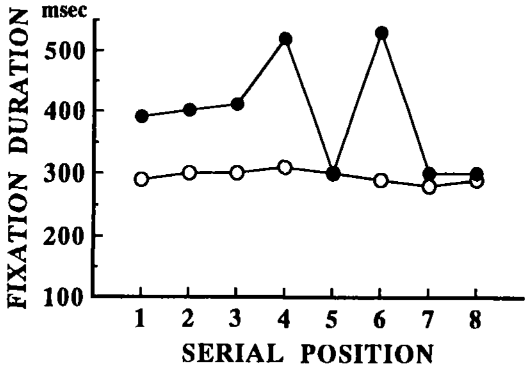
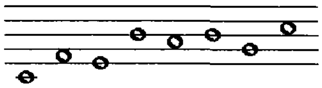
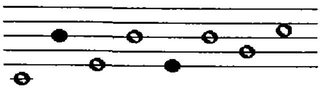
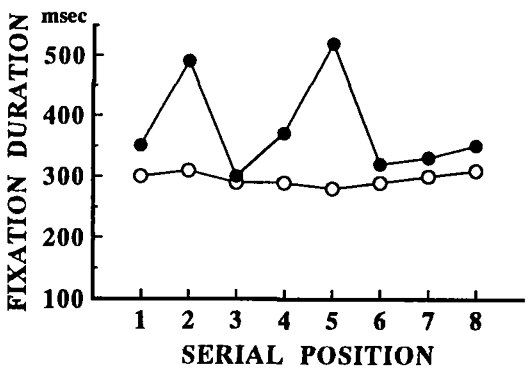
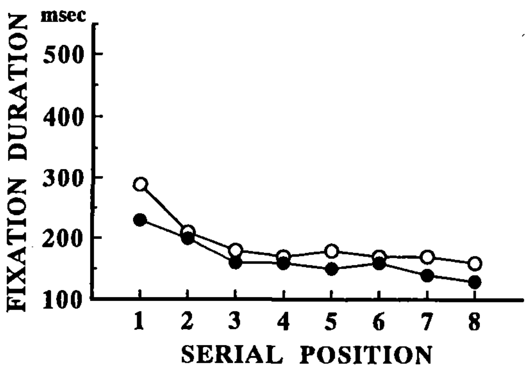
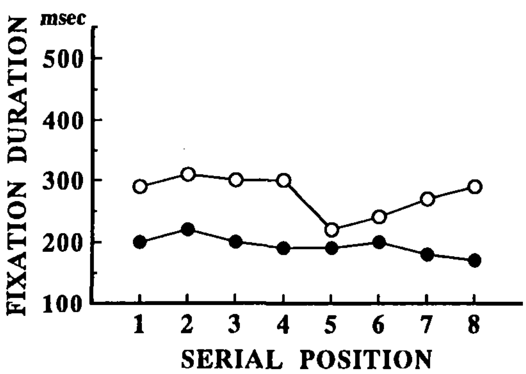


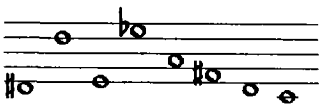
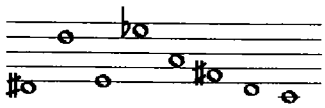
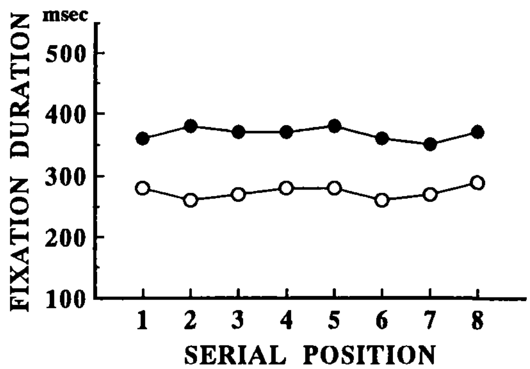
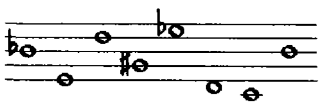
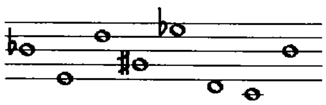
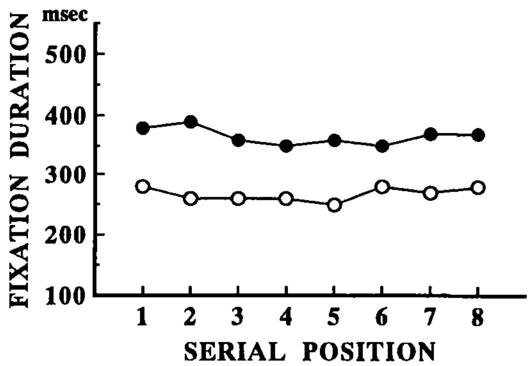
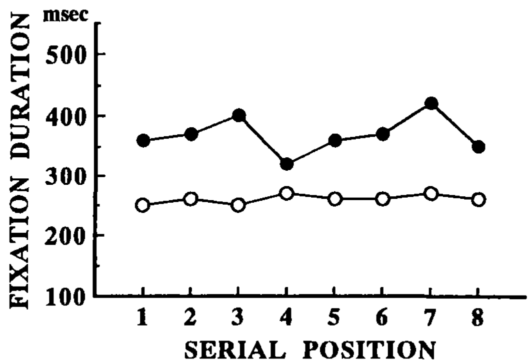
Figure 5–1.

Average duration of fixation for Groups H and L as a function of the serial position in the latter half of each pattern in the Staff Notation condition. In the visual tracking task, black notes indicate that their positions are deviated from the exact points.

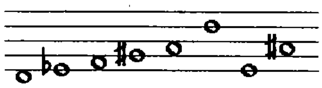
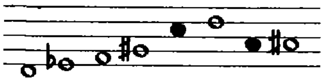
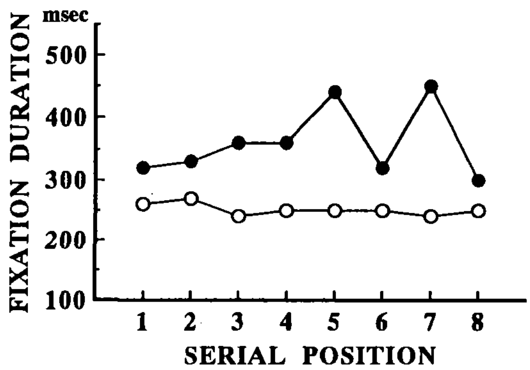
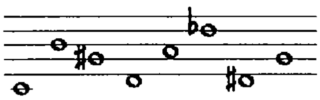
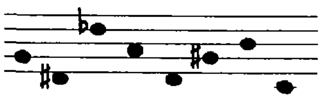
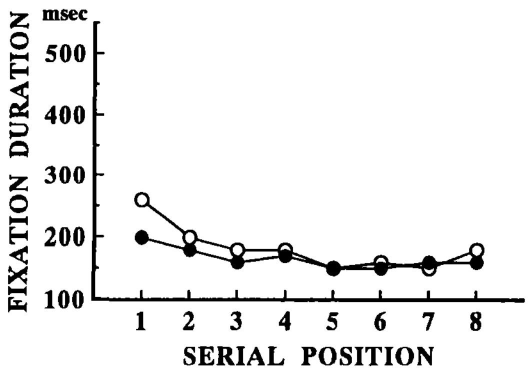
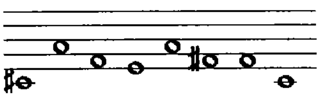
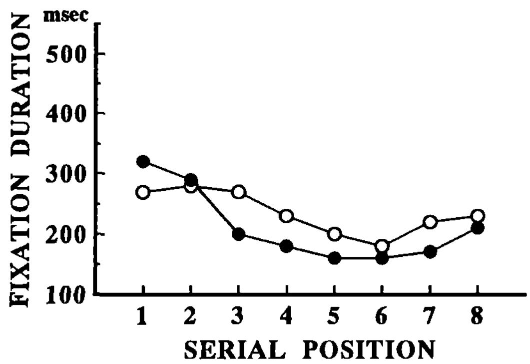
Staff Notation

	Standard Melody	Visual Tracking Task	Fixation Duration (● Group H, ○ Group L)
3		Tonal C 	
4		Tonal C 	
5		Tonal R 	
6		Tonal R 	

Staff Notation

	Standard Melody	Visual Tracking Task	Fixation Duration ● Group H ○ Group L
7		Atonal S 	
8		Atonal S 	
9		Atonal C 	

Staff Notation

	Standard Melody	Visual Tracking Task	Fixation Duration ● Group H ○ Group L
10		Atonal C 	
11		Atonal R 	
12		Atonal R 	

Melodic Contour

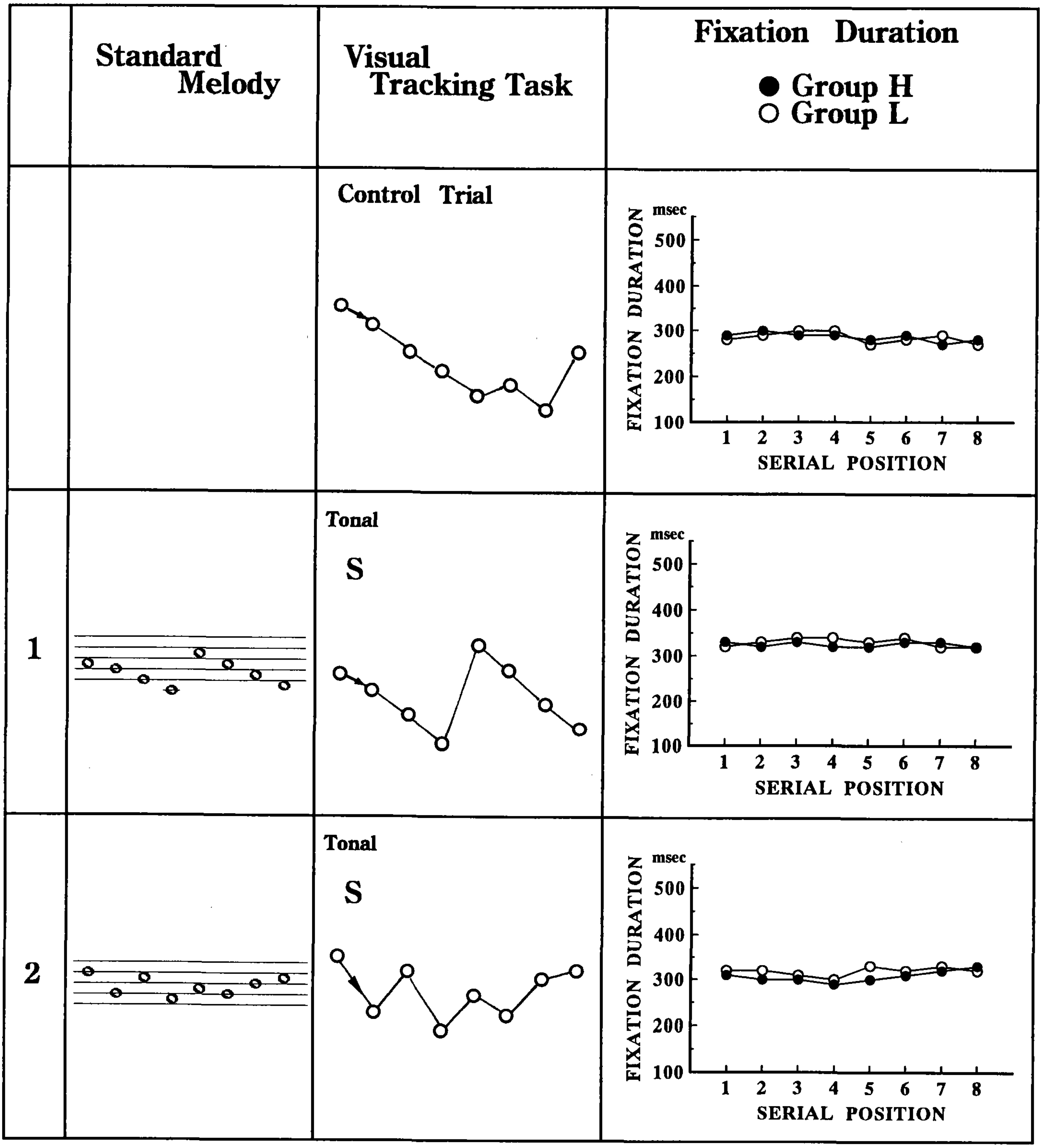
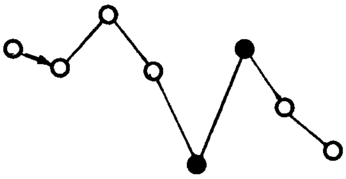
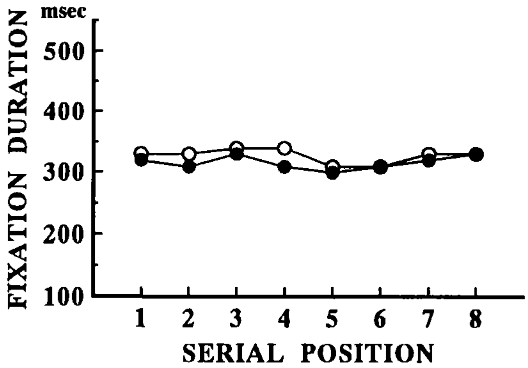
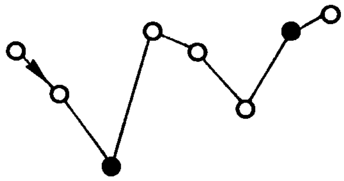
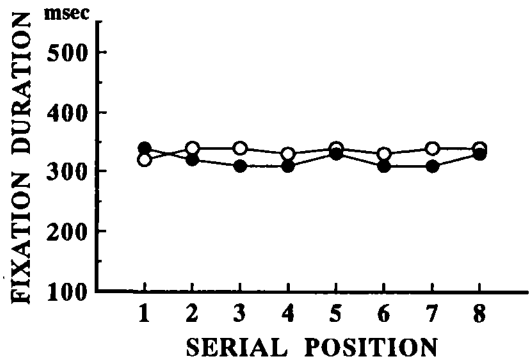
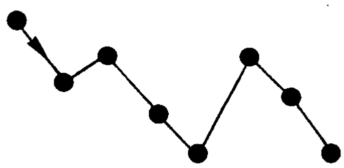
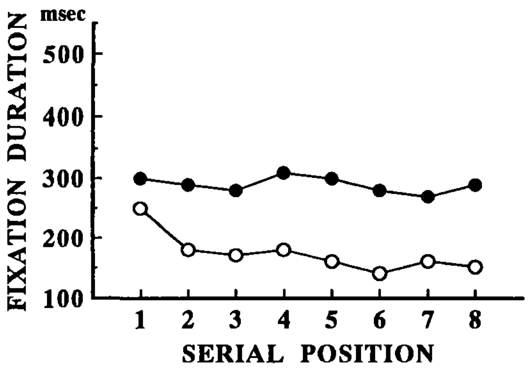
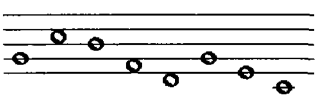
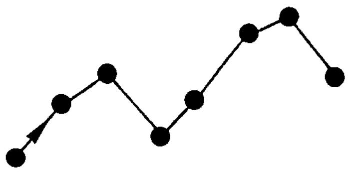
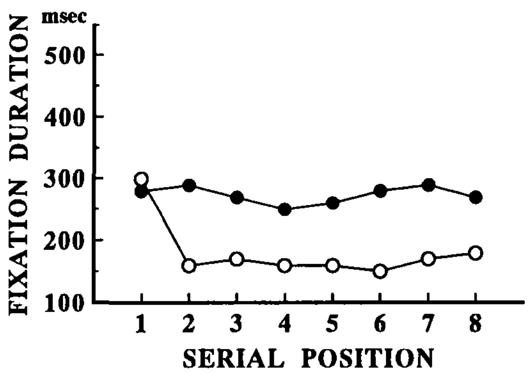


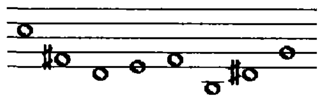
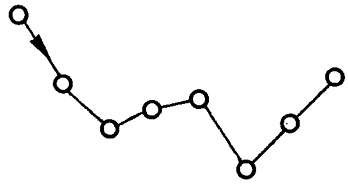
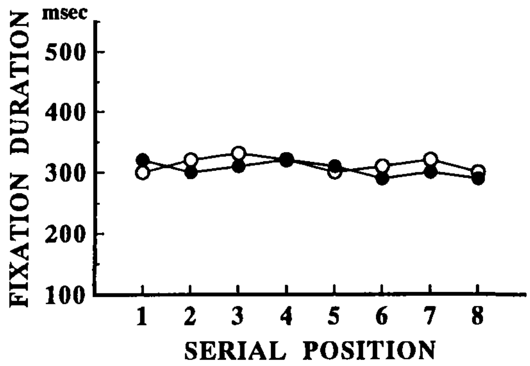
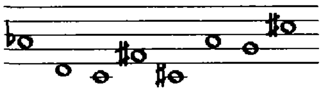
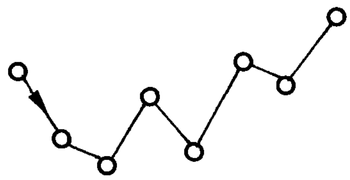
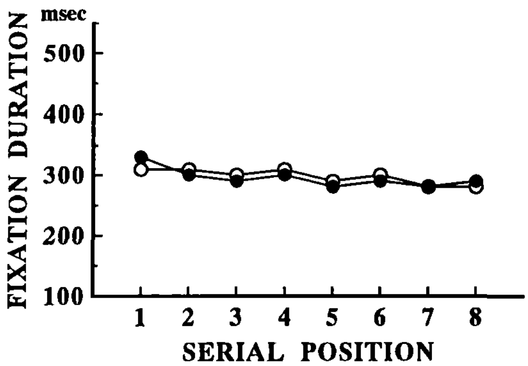
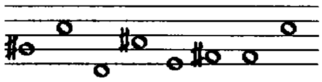
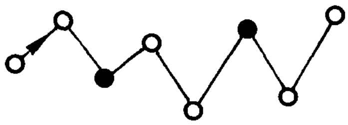
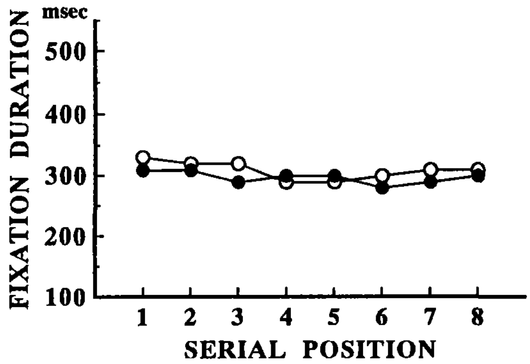
Figure 5-2.

Average duration of fixation for Groups H and L as a function of the serial position in the latter half of each pattern in the Melodic Contour condition. In the visual tracking task, black circles indicate that their positions are deviated from the exact points.

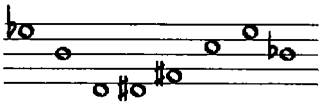
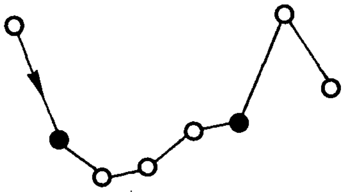
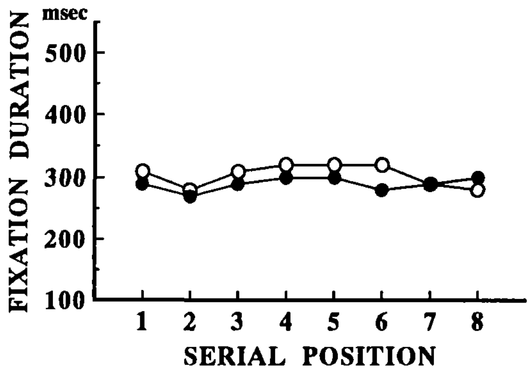
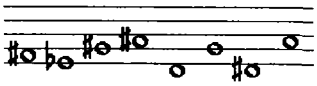
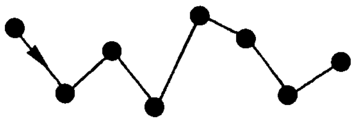
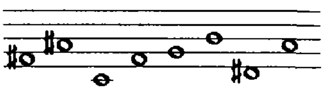
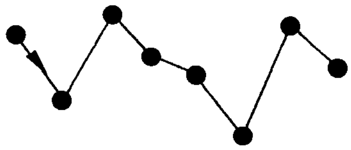
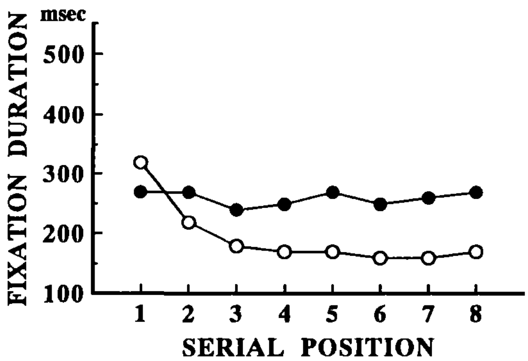
Melodic Contour

	Standard Melody	Visual Tracking Task	Fixation Duration ● Group H ○ Group L
3		Tonal C 	
4		Tonal C 	
5		Tonal R 	
6		Tonal R 	

Melodic Contour

	Standard Melody	Visual Tracking Task	Fixation Duration ● Group H ○ Group L
7		Atonal S 	
8		Atonal S 	
9		Atonal C 	

Melodic Contour

	Standard Melody	Visual Tracking Task	Fixation Duration ● Group H ○ Group L
10		Atonal C 	
11		Atonal R 	
12		Atonal R 	

Staff Notation

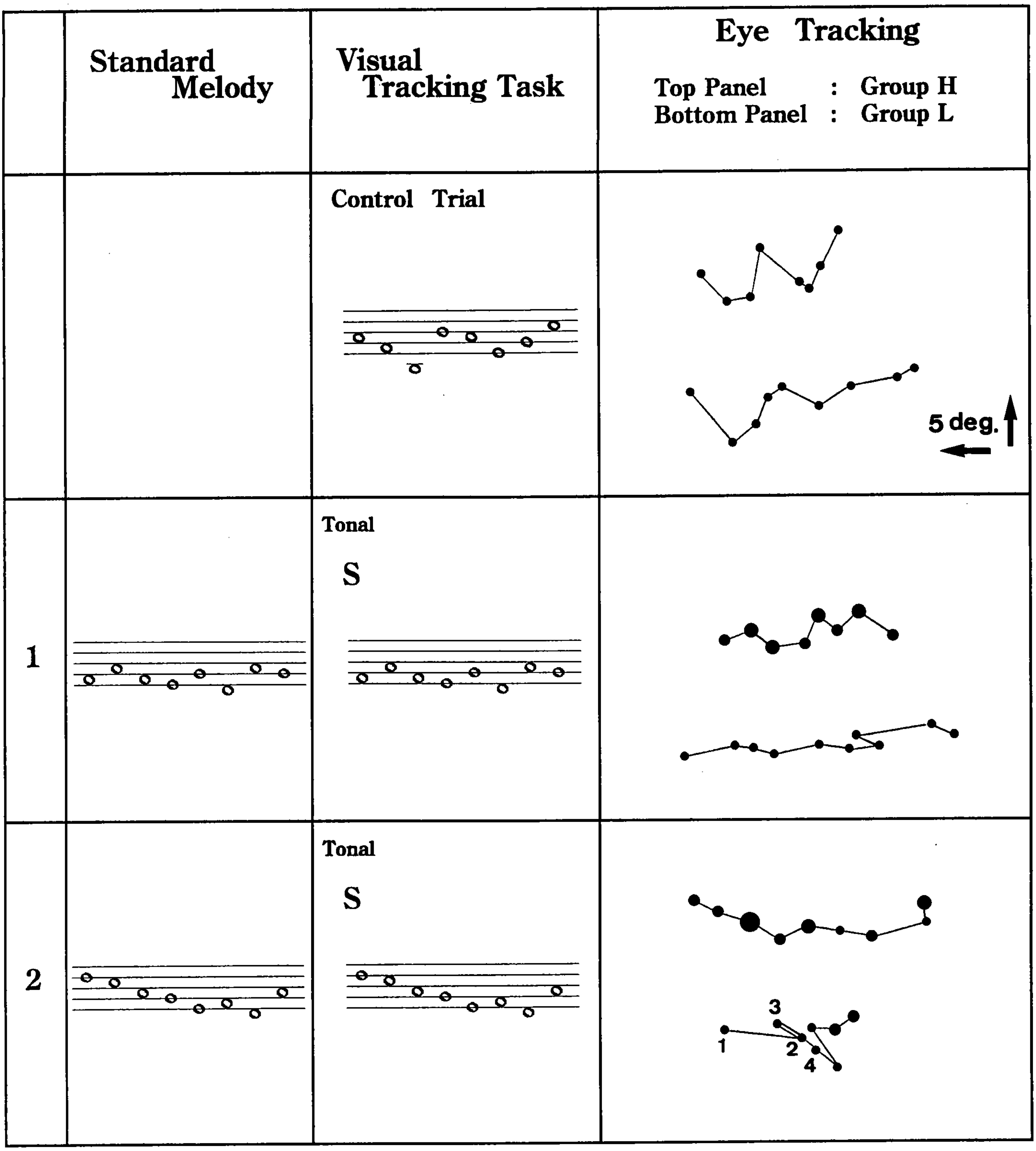
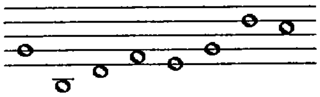
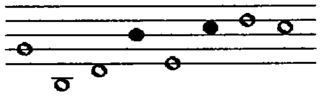
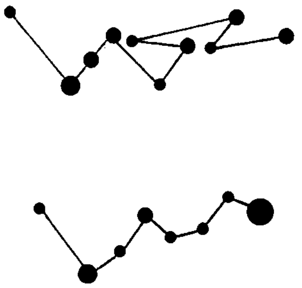
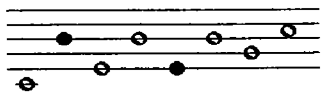
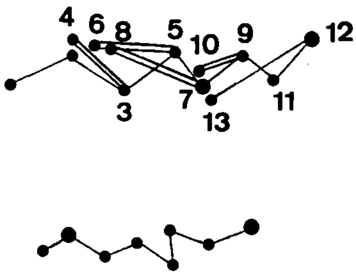
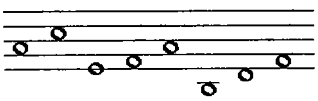
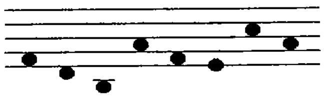
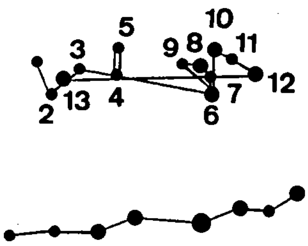
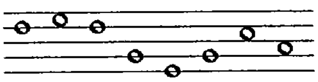
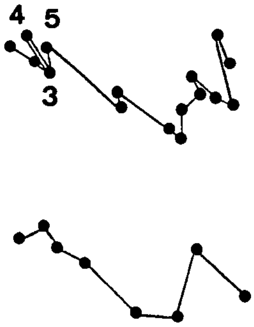


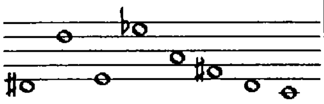
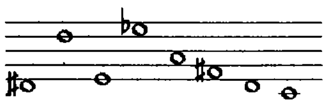
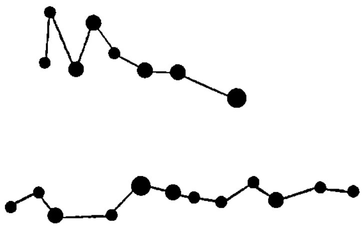
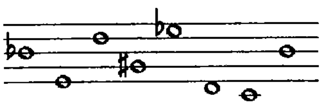
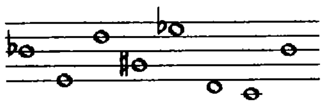
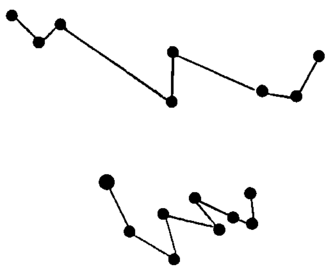
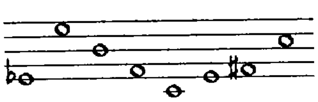
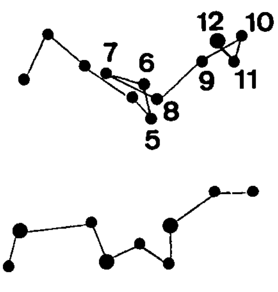
Figure 5-3.

Samples of eye tracking data in the latter half of each pattern for subjects in Groups H and L in the Staff Notation condition. Fixation durations of more than 100 msec with visual angles of less than 2.0 degree are marked by circles, and larger circles indicate that subjects fixated on the points for a longer duration. Sequential fixations are connected by straight lines. Fixations are numbered to show the order of occurrence on the scanpath. In the visual tracking task, black notes indicate that their positions are deviated from the exact points.

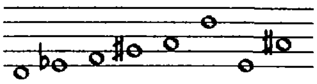
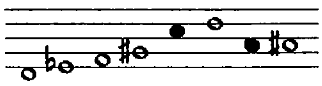
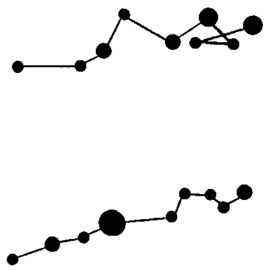
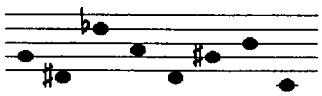
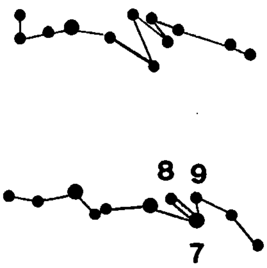
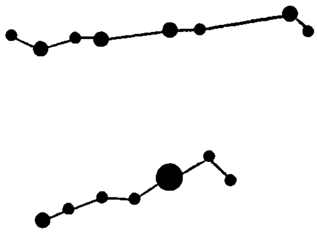
Staff Notation

	Standard Melody	Visual Tracking Task	Eye Tracking Top Panel : Group H Bottom Panel : Group L
3		Tonal C 	
4		Tonal C 	
5		Tonal R 	
6		Tonal R 	

Staff Notation

	Standard Melody	Visual Tracking Task	Eye Tracking Top Panel : Group H Bottom Panel : Group L
7		Atonal S 	
8		Atonal S 	
9		Atonal C 	

Staff Notation

	Standard Melody	Visual Tracking Task	Eye Tracking Top Panel : Group H Bottom Panel : Group L
10		Atonal C 	
11		Atonal R 	
12		Atonal R 	

Melodic Contour

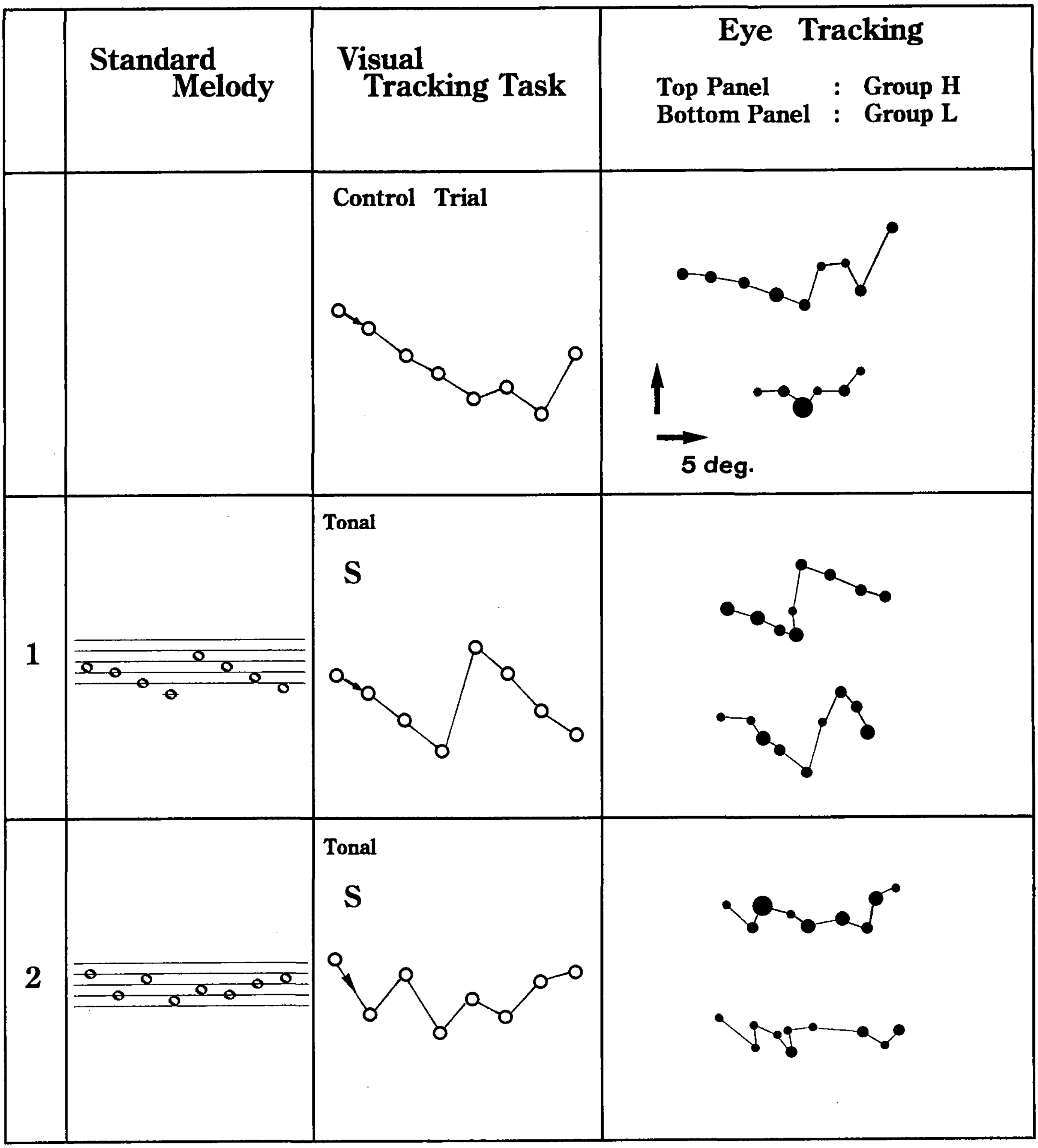
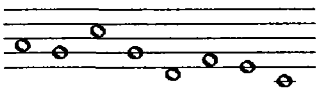
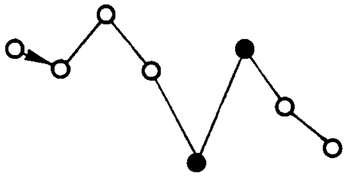
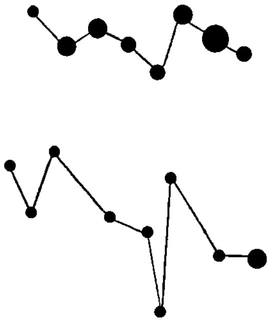
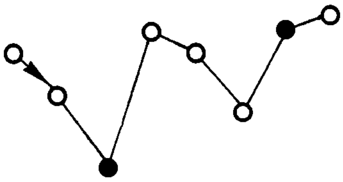
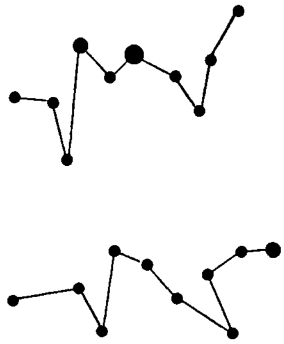
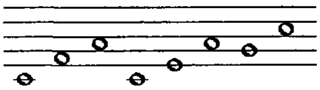
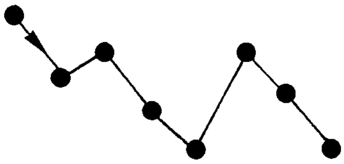
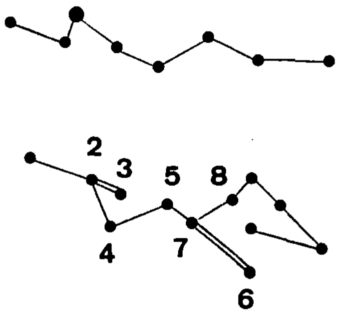
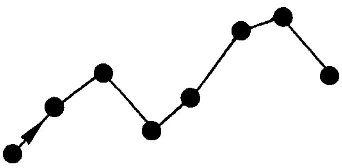
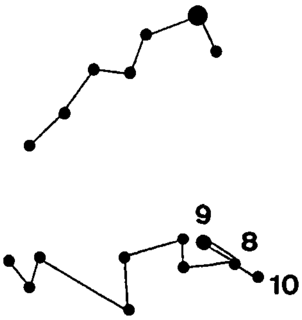


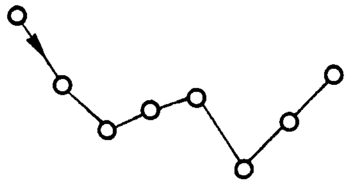
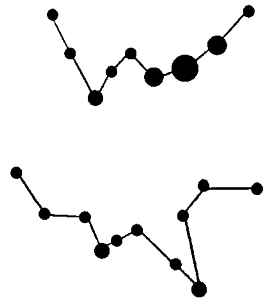
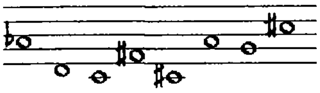
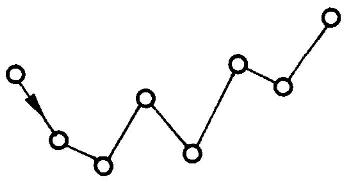
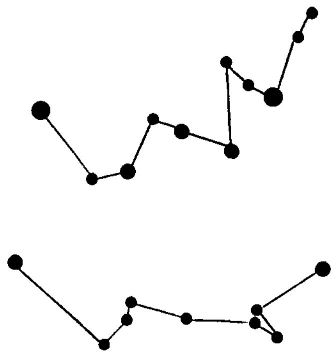
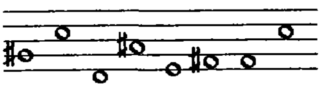
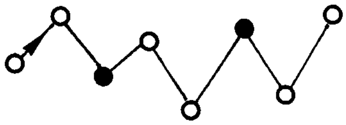
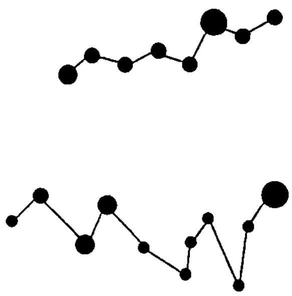
Figure 5-4.

Samples of eye tracking data in the latter half of each pattern for subjects in Groups H and L in the Melodic Contour condition. Fixation durations of more than 100 msec with visual angles of less than are marked by circles, and larger circles indicate that subjects fixated on the points for a longer duration. Sequential fixations are connected by straight lines. Fixations are numbered to show the order of occurrence on the scanpath. In the visual tracking task, black circles indicate that their positions are deviated from the exact points.

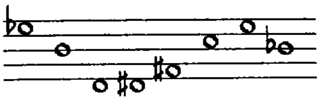
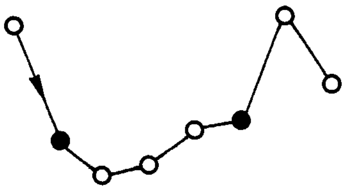
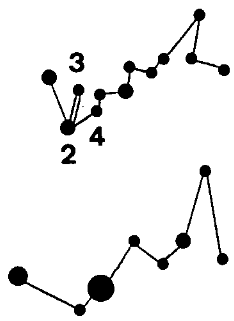
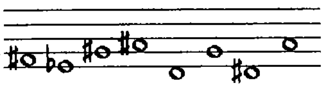
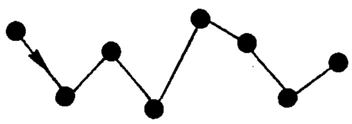
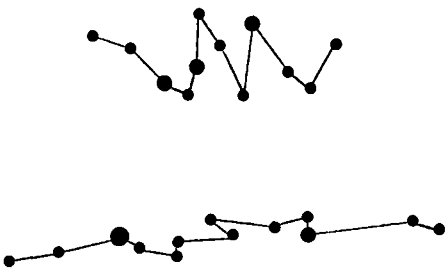
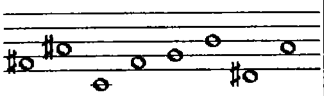
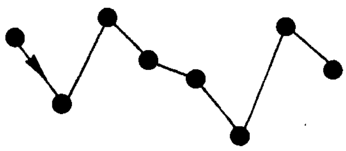
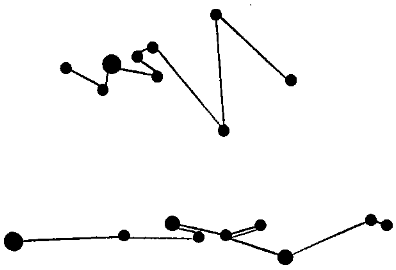
Melodic Contour

	Standard Melody	Visual Tracking Task	Eye Tracking Top Panel : Group H Bottom Panel : Group L
3		Tonal C 	
4		Tonal C 	
5		Tonal R 	
6		Tonal R 	

Melodic Contour

	Standard Melody	Visual Tracking Task	Eye Tracking Top Panel : Group H Bottom Panel : Group L
7		Atonal S 	
8		Atonal S 	
9		Atonal C 	

Melodic Contour

	Standard Melody	Visual Tracking Task	Eye Tracking Top Panel : Group H Bottom Panel : Group L
10		Atonal C 	
11		Atonal R 	
12		Atonal R 	

CHAPTER 4

EXPERIMENTAL TEST OF MULTI ENCODING STRATEGIES FOR PITCH

EXPERIMENT 6

Acoustic, verbal, and visual codes

EXPERIMENT 6

Acoustic, Verbal, and Visual Codes

In Experiment 1, highly musically trained subjects reported that they encode pitches of atonal melodies as acoustic pitch codes by humming or mental rehearsal of pitches, where–as they encode pitches of tonal melodies as note names accompanied by corresponding pitches rather than as merely note names in a monotone. Thus, it is quite possible that they employ a dual coding strategy to encode pitches of tonal melodies. Furthermore, in Exper–iment 4, it was observed that some subjects rehearsed pitches or note names along with the pitches of the standard melody during the visual tracking task. They reported that they often visualized pitch information as melodic contours or staff notations, which were almost always accompanied by pitch rehearsal or verbal rehearsal of note names along with pitches. These findings suggest the further possibility that when subjects employed a visual–izing strategy it was a dual or triple coding strategy, such as melodic contour accompanied by pitch rehearsal (dual) or staff notation accompanied by pitch rehearsal with note names (triple).

Bennett, Hausfeld, Reeve, and Smith (1978) reported that musicians may be able to use any or all of pitch, naming or visual codes, depending on the situation at any time : thus, when one code is assessed as likely to be ineffective, another alternate process is switched in. Furthermore, subjects would be motivated to attempt to engage in covert rehearsal (Muter, 1980) with multi–code sets.

In Experiment 6, some multi-distractor sets were prepared to interfere with the operation of such multi-code sets. The type of encoding strategy employed during the retention interval was inferred from the disruptive effects of the multi-distractor sets on the recognition performance. The inference was made under the assumption that a distractor which has high similarity to the encoding strategy would have considerable disruptive effect on the recognition performance.

Method

Subjects

Thirty-three female subjects (average age, 20.2 years ; age range, 18–22 years) who were undergraduates majoring in music constituted the highly musically trained group (Group H). Each had had at least 12 years of formal training in playing the piano, with an average of 14.8 years (range, 12–19 years).

Fifty-two female subjects (average age, 20.3 years ; age range, 18–22 years) who were undergraduates majoring in literature, education, or domestic science constituted the less well musically trained group (Group L). Their musical experience on an instrument averaged 4.2 years (range, 0–6 years).

Materials

Each trial involved a standard series followed by a retention interval and then by a comparison series. Both the standard and comparison series consisted of eight tones taken from an equal-temperament scale, which ranged from A_3 (220.0 Hz) to E_5 (659.26 Hz). The duration of each tone was 700 msec (600 msec with 100-msec silence), so that the length of the 8-tone series was 5.6 sec.

For this experiment, three lists were prepared, and each list involved 96 trials. In half of them (48 trials), tonal melodies in a major key which were high in tonal melodic structure according to conventional Western rules were used. In the other half (48 trials), atonal melodies which were low in tonal melodic structure were used. Both types of melodies involved a wide variety of contours. Half of the comparison series were exactly the same as the standard series, and the other half were different from the standard.

Twelve types of interfering stimuli were interpolated during the retention interval between the standard and the comparison series. "Pause" (**P**) was a blank retention interval. "Series of musical Note Names" (**NN**) consisted of a series of eight words chosen randomly from seven note names (Do, Re, Mi, Fa, Sol, La, Si). The series were delivered in a monotonous pitch in a female voice. "Interfering Melody" (**IM**) was composed of eight tones taken from an equal-temperament scale, which ranged from A_3 (220.0 Hz) to E_5 (659.26 Hz) having either a tonal structure (for a tonal trial) or an atonal structure (for an atonal trial). "Staff Notation" (**SN**) consisted of a series of eight musical notes ranged from A_3 to E_5 on a staff without a G clef, and the eight notes constituted tonal or atonal melody. On a computer display, the eight notes appeared successively from the left to the right with a constant space interval of about 2.5 cm on the staff (line interval : 2.0 cm). "Melodic Contour" (**MC**) was an ascending and descending pattern of melody, and was represented as eight circles with a diameter of 0.5 cm connected by seven line segments, which also appeared successively from the left to the right on the computer display. "Interfering Melody + Note Names" (**IM+NN**) was an 8–tone interfering melody sung with note names at accurate pitches in a female voice. "Interfering Melody + Staff Notation" (**IM+SN**) was an 8–tone interfering melody presented with staff notation ; each tone of a melody was presented synchronously with each note appearing on the staff, with their pitches coinciding exactly. "Interfering Melody + Melodic Contour" (**IM+MC**) was an 8–tone interfering melody presented with melodic contour ; each tone of a melody was presented synchronously with each circle in the melodic contour, with their pitch intervals coinciding exactly. "Note Names + Staff Notation" (NN+SN) was a series of eight note names presented with staff notation ; each note name was presented synchronously with each note appearing on the staff, with their pitches coinciding exactly. "Note Names + Melodic Contour" (NN+MC) was a series of eight note names presented with melodic contour ; each note name was presented synchronously with each circle in the melodic contour, with their pitch intervals coinciding exactly. " Interfering Melody + Note Names + Staff Notation" (IM+NN+SN) was an 8-tone interfering melody sung with note names at accurate pitches and each tone of the melody was presented synchronously with each note appearing on the staff, with their pitches coinciding exactly. "Interfering Melody + Note Names + Melodic Contour" (IM+NN+MC) was an 8-tone interfering melody sung with note names at accurate pitches ; each tone of the melody was presented synchronously with each circle in the melodic contour, with their pitch intervals coinciding exactly.

Thus, NN, IM, and IM+NN were auditory stimuli, SN and MC were visual stimuli, and IM+SN, IM+MC, NN+SN, NN+MC, IM+NN+SN and IM+NN+MC were auditory-visual combination stimuli.

The series of Note Names (NN) was constructed as follows : Seven note names (Do, Re, Mi, Fa, Sol, La, Si) delivered with monotonous pitch in a professional female voice were recorded, and each of them was then cut to 600 msec with sampling frequency of 32 KHz. Then, a serial arrangement consisting of eight note names separated by 100 msec was constructed. If the NN was presented together with another interfering stimulus, it was arranged to correspond to the other interfering stimulus.

Interfering Melody with Note Names (IM+NN) was constructed as follows : Twelve pitches sung with note names at accurate pitches in a professional female voice were recorded (A_3=220.00, B_3=246.94, C_4=261.63, D_4=293.66, E_4=329.63, F_4=349.23, G_4=392.00, A_4=440.00, B_4=493.88, C_5=523.25, D_5=587.33, E_5=659.26). Each of them was then cut to 600 msec with sampling frequency of 32 KHz. Then, a serial arrangement consisting of eight pitches separated by 100 msec was constructed for either a tonal or an atonal structure. If IM+NN was presented together with another interfering stimulus, it was arranged to correspond to the other interfering stimulus. These operations were performed using a CANOPUS Sound Master/V.

The duration of each tone in an auditory stimulus was 700 msec (600 msec with 100-msec silence), and the notes or circles in a visual stimulus appeared successively every 700 msec, and each series (700 msec X 8 = 5.6 sec) was repeated twice during the retention interval. Therefore, each trial involved a 5.6-sec standard series followed by a 2.0-sec interval, an 11.2-sec interfering stimulus, a 2.0-sec interval, and then a 5.6-sec comparison series.

The tones of the standard, comparison and interfering melodies were generated by an NEC PC-9801 26K sound synthesizer board installed in an NEC PC-9801 DA personal computer, recorded on tape, and presented over high-quality sound reproduction equipment (DENON digital audio tape recorder DTR-80P).

Procedure

Each subject sat in front of a personal computer and wore headphones, and all tones were adjusted to be equal in loudness. The subjects were instructed that this was an experiment on memory for melodies, and that in each of the 96 trials, they would first hear a trial number, then a first melody (standard series), which would be followed by an interfering stimulus, and then a second melody (comparison series). The subject's task was to judge whether the two melodies were the same or different in pitch, and to indicate their judgments by writing "S" (Same) or "D" (Different) on an answer sheet. The subjects were also instructed that they should track the visual stimuli, but that they should not employ any motor encoding strategy analogous to playing the piano. Three practice trials with feedback were given to the subjects prior to the 96 trials, and the interval between trials was 10 sec.

In this experiment, there were 48 tonal trials and 48 atonal trials. In both types of melodies, 4 trials were run in each type of interference condition. The 12 types of interference conditions were alternated, and the order of presentation of all 96 trials was randomized. As described above, for this experiment, three lists were prepared, and each list involved 96 trials. These lists were counterbalanced between subjects, to minimize the differences among the trials in each of the 12 interference conditions in difficulty.

Results and Discussion

Recognition probability data (hit rate plus correct rejection rate) for Groups H and L were analyzed in a two-way analysis of variance [2 Melody Types X 12 Interference Conditions], with repeated measures on both factors. For Group H, there were significant main effects of Melody Type and Interference Condition [$F(1,32)=149.86, p<.001$; $F(11,352)=9.06, p<.001$], and there was a significant interaction of Melody Type X Interference Condition [$F(11,352)=2.51, p<.001$]. For Group L, there were significant main effects of Melody Type and Interference Condition [$F(1,51)=86.08, p<.001$; $F(11,561)=6.74, p<.001$]. For both Groups H and L, the main effect of Melody Type indicates that the subjects performed better with tonal than with atonal melodies, and the main effect of Interference Condition indicates that there were significant disruptive effects in some of these 12 interference conditions. The interaction of Melody Type X Interference Condition in Group H indicates that there were differences in the disruptive effects of the interference conditions on the performance between the two melody types. There was no significant interaction of Melody Type X Interference Condition in Group L, therefore, the disruptive effects of the interference conditions on the performance with the tonal melodies were the same as those on that with the atonal melodies.

Recognition performance in Group H for the tonal melodies is shown in Figure 6–1. In the auditory interference conditions (Figure 6–1(a)), the recognition probability in IM was lower than that in P ($p<.05$ by Newman Keuls method), that in NN was lower than that in P ($p<.01$), and that in IM+NN was lower than those in P, IM (both $p<.01$), and NN ($p<.05$). In the auditory–visual MC combination interference conditions (Figure 6–1(b)), the recognition probability in MC was somewhat but not significantly lower than that in P, while those in NN+MC and IM+MC were lower than that in P (both $p<.01$), and that in IM+NN+MC was lower than those in P, MC (both $p<.01$), and NN+MC ($p<.05$). In the auditory–visual SN combination interference conditions (Figure 6–1(c)), the recognition probability in SN was lower than that in P ($p<.05$), those in IM+SN and NN+SN were lower than that in P (both $p<.01$), and that in IM+NN+SN was lower than those in P, SN (both $p<.01$), and IM+SN ($p<.05$). Thus, in comparison with that in P, the recognition probabilities in all interference conditions except for MC were significantly lower, and those in IM+NN and IM+NN+SN were the lowest of all conditions. The significantly lower recognition probability in the interference conditions compared to that in the control condition P indicated that the recognition performance was significantly disrupted by the interference condition, and it could be concluded that the subjects might have used the same strategy as that incorporated in the interfering stimulus to encode pitches of the standard melodies.

These results raise the following suggestions : For the subjects of Group H, visualizing melodic contours is not an effective strategy to encode pitches of tonal melodies. As was also found in Experiment 1, the verbal rehearsal of note names is somewhat more effectual than pitch rehearsal strategy through such processes as humming, whistling, or mental rehearsal of pitches. Visualizing staff notation is also somewhat effectual. In fact, however, the subjects used dual or triple coding strategies rather than such single coding strategies. Pitch rehearsal of auditory information along with note names (dual coding) or, to an even greater extent, at the same time visualizing the staff notation (triple coding) were the most effectual strategies by which the Group H subjects memorized and retained pitches of tonal melodies.

Recognition performance in Group H for the atonal melodies is shown in Figure 6–2. In the auditory interference conditions (Figure 6–2(a)), the recognition probabilities in NN and IM were somewhat but not significantly lower than that in P, while that in IM+NN was lower than that in P ($p<.05$). In the auditory–visual MC combination interference conditions (Figure 6–2(b)), the recognition probability in NN+MC was somewhat but not significantly lower than that in P, while that in MC was lower than that in P ($p<.05$), and those in IM+MC and IM+NN+MC were lower than that in P (both $p<.01$). In the auditory–visual SN combination interference conditions (Figure 6–2(c)), the recognition probabilities in SN and NN+SN were somewhat but not significantly lower than that in P, while those in IM+SN and IM+NN+SN were lower than that in P (both $p<.05$).

These results raise the following suggestions : For the subjects in Group H, a single encoding strategy, such as verbal rehearsal of note names, pitch rehearsal of auditory information, or visualizing staff notation, is not an effective strategy to encode pitches in atonal melodies. As was also found in Experiment 1, the results of this experiment suggest that it was difficult for the subjects to encode pitches in atonal structure as musical note names. It was also difficult for them to accurately visualize notes on a staff, unless the pitches were identified as their note names. Although it was found in Experiment 1 that the subjects used an auditory encoding strategy in which pitches in an atonal melody were retained in memory as auditory information, in this experiment the auditory encoding strategy was found to be ineffectual. This discrepancy is considered to be due to a difference in the standard melody, which in Experiment 1 consisted of 6 tones and their pitches could be encoded to some extent by pitch maintenance rehearsal of auditory information, but which in this experiment consisted of 8 tones and their pitches might have been difficult to encode by such pitch maintenance rehearsal. The subjects would have attempted to use pitch rehearsal along with note names, which was found to be the most effectual strategy for tonal melodies, but would have found it difficult to encode pitches in atonal structure accurately by this strategy. Therefore, encoding of the pitches of atonal melodies depends considerably on visualization of the melodic contour or an auditory–visual combination strategy by pitch rehearsal along with visualization of the melodic contour. The recognition performance was disrupted to some extent by the auditory–visual SN combination interference conditions, because the subjects attempted to visualize the staff notation of even pitches in atonal structure or would abstract melodic contours from note sequences on a staff. Multi coding strategies obtained in Group H with both tonal and atonal melodies support the findings that melodic intervals are encoded relatively well with tonal melodies (Bartlett & Dowling, 1980), and that in short–term memory, contour is important when the tonal context is week or confusing and not aided by meaningful tonal context (Dowling, 1982).

Recognition performance in Group L for the tonal melodies is shown in Figure 6–3. In the auditory interference conditions (Figure 6–3(a)), the recognition probabilities in NN, IM and IM+NN were somewhat but not significantly lower than that in P. In the auditory–

visual MC combination interference conditions (Figure 6–3(b)), the recognition probabilities in MC and NN+MC were somewhat but not significantly lower than that in P, while that in IM+MC was lower than that in P (p<.05), and that in IM+NN+MC was lower than those in P, MC (both p<.01), and NN+MC (p<.05). In the auditory–visual SN combination interference conditions (Figure 6–3(c)), the recognition probabilities in SN, NN+SN, IM+SN and IM+NN+SN were somewhat but not significantly lower than that in P.

Recognition performance in Group L for the atonal melodies is shown in Figure 6–4. In the auditory interference conditions (Figure 6–4(a)), the recognition probabilities in NN, IM and IM+NN were somewhat but not significantly lower than that in P. In the auditory–visual MC combination interference conditions (Figure 6–4(b)), the recognition probability in MC was somewhat but not significantly lower than that in P, while that in NN+MC was lower than that in P (p<.05), and those in IM+MC and IM+NN+MC were lower than that in P (both p<.01). In the auditory–visual SN combination interference conditions (Figure 6–4(c)), the recognition probabilities in SN and NN+SN were somewhat but not significantly lower than that in P, while those in IM+SN and IM+NN+SN were lower than that in P (both p<.05).

As described above, there was no interaction of Melody Type X Interference Condition in Group L, therefore the disruptive effects of interference conditions on the performance with the tonal melodies were the same as those with the atonal melodies. Moreover, these results were similar to that in Group H with the atonal melodies, although the total recognition performance in Group L was significantly lower than that in Group H (p<.001). The subjects in Group L did not use a single encoding strategy, such as verbal rehearsal of note names, pitch rehearsal of auditory information, visualizing staff notation, or even visualizing melodic contour. The encoding of the pitches of melodies would depend con–

siderably on auditory–visual combination strategy by pitch rehearsal along with visualization of the melodic contour. The greatest difference between the performance in Group L with both types of melodies and that in Group H with atonal melodies was the significant disruption by the interference condition IM+NN in the latter, in contrast to the lack of disruption in the former. Thus, the subjects in Group H attempted to some extent to encode pitches of atonal melodies by pitch rehearsal of note names, but it was very difficult for the subjects in Group L to use this strategy.

These findings suggest that the subjects encoded pitches of short melodies using two or three codes rather than just one. For example, pitch rehearsal of auditory information along with note names (dual coding) and, to an even greater extent, at the same time visualizing the staff notation (triple coding) were the most effective strategies for Group H with pitches of tonal melodies ; pitch rehearsal along with visualization of melodic contour (dual coding) was also effective for Group H with atonal melodies and for Group L with both types of melodies.

As described in Chapter 1, Paivio proposed the dual–coding theory (Paivio, 1971, 1978), the essence of which is that there are two basic independent but interconnected systems for the representation and processing of information. The verbal system deals with linguistic information and stores it in an appropriate verbal form, while the nonverbal system carries out image–based processing and representation.

The dual–coding by pitch rehearsal along with note names found in this experiment is consistent with Paivio's theory, because pitch rehearsal would be in the nonverbal system and note names would be in the verbal system. If the triple–coding by visualizing staff notation along with pitch rehearsal with note names is consistent with the Paivio's theory, the complexity of the explanation of such triple codes by Paivio's dual–coding theory suggests that this theory might be somewhat oversimplified.

However, Paivio (1986) proposed that within the two systems there are associative structures. Both systems have basic representational units that are linked to one another by referential connections : Logogens for the verbal system and Imagens for the nonverbal system. Both the Logogens and Imagens are further divided into sub-systems (i.e., visual, auditory, kinesthetic, gustatory, and olfactory) which process either verbal or nonverbal information in the different modalities (modality-specific).

Therefore, the dual-coding by pitch rehearsal along with note names found in this experiment is consistent with Paivio's theory again, because pitch rehearsal would be an auditory Imagen and note names would be a auditory Logogen. If the triple-coding by visualizing staff notation along with pitch rehearsal with note names is consistent with the Paivio's theory, staff notation must be a visual Imagen, pitch rehearsal must be an auditory Imagen, and note names must be a auditory Logogen. Furthermore, the dual-coding by visualizing melodic contour along with pitch rehearsal would be considered the operation of a single system in Paivio's theory, because the melodic contour would be a visual Imagen and pitch rehearsal would be an auditory Imagen, and both codes would be together in a nonverbal system. The findings of Experiment 6 are interpreted using Baddeley's working memory theory in Chapter 7.

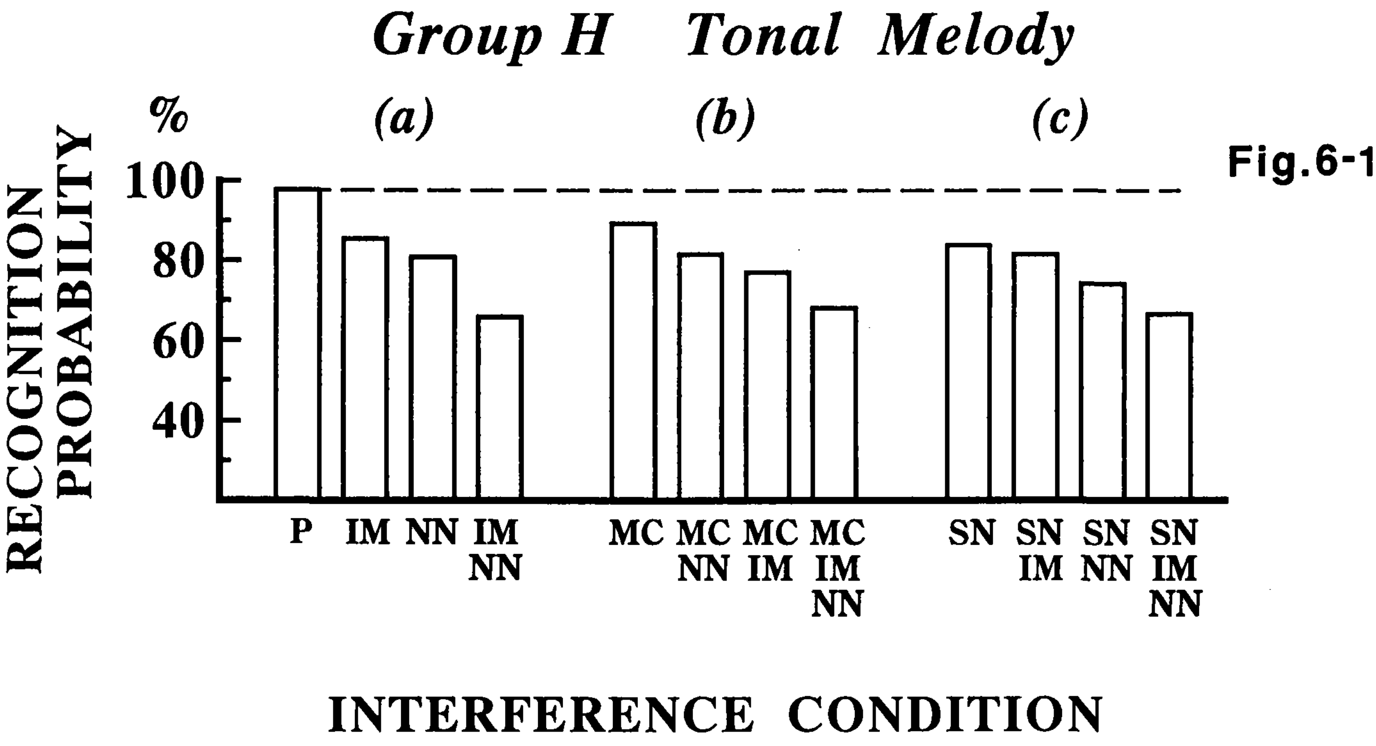


Figure 6.

Mean probability of correct recognition (hit rate plus correct rejection rate) for the 12 interference conditions for tonal and atonal melodies in Groups H and L.

(a) auditory interference conditions
(b) auditory–visual MC combination interference conditions
(c) auditory–visual SN combination interference conditions

Notes.

(a)
P : Pause,
IM : Interfering Melody,
NN : series of musical Note Names,
IM+NN : Interfering Melody + Note Names,

(b)
MC : Melodic Contour,
IM+MC : Interfering Melody + Melodic Contour,
NN+MC : Note Names + Melodic Contour,
IM+NN+MC : Interfering Melody + Note Names + Melodic Contour.

(c)
SN : Staff Notation,
IM+SN : Interfering Melody + Staff Notation,
NN+SN : Note Names + Staff Notation,
IM+NN+SN : Interfering Melody + Note Names + Staff Notation.

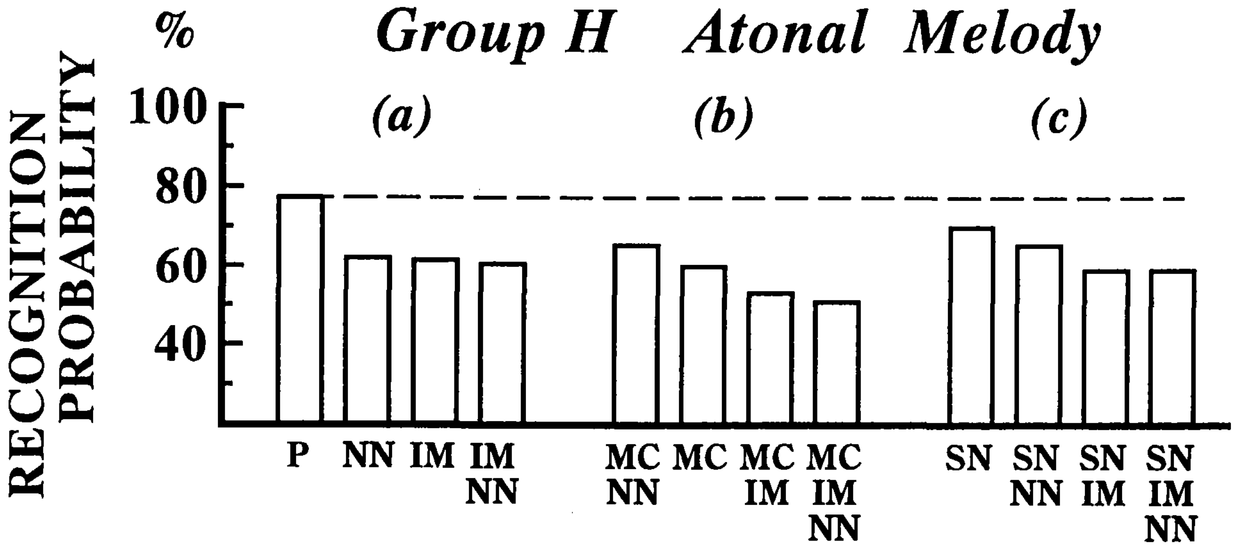


Fig.6-2

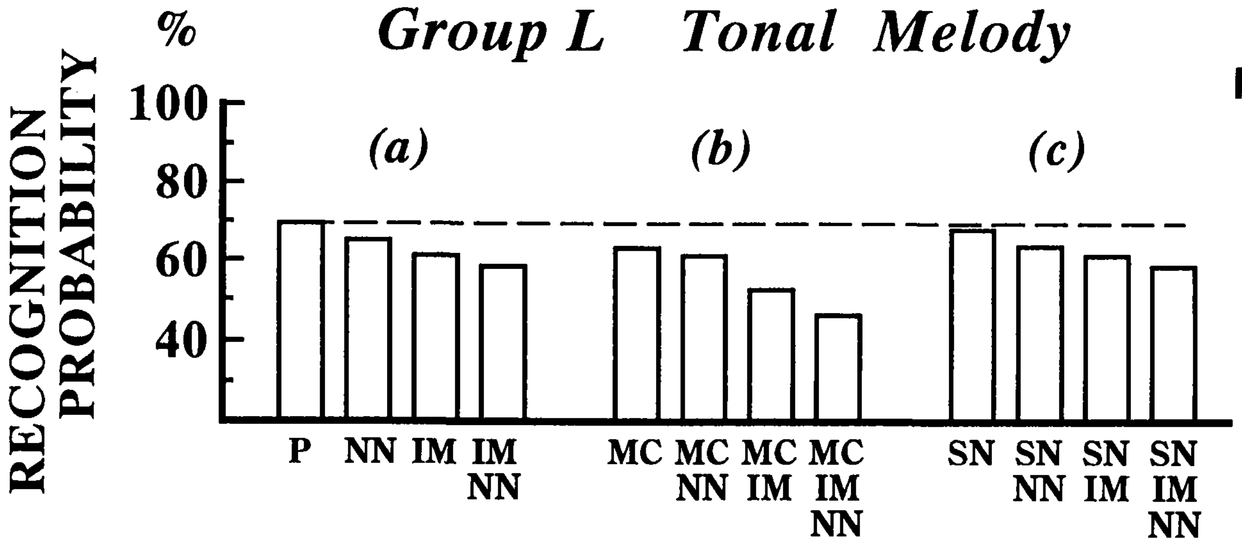


Fig.6-3

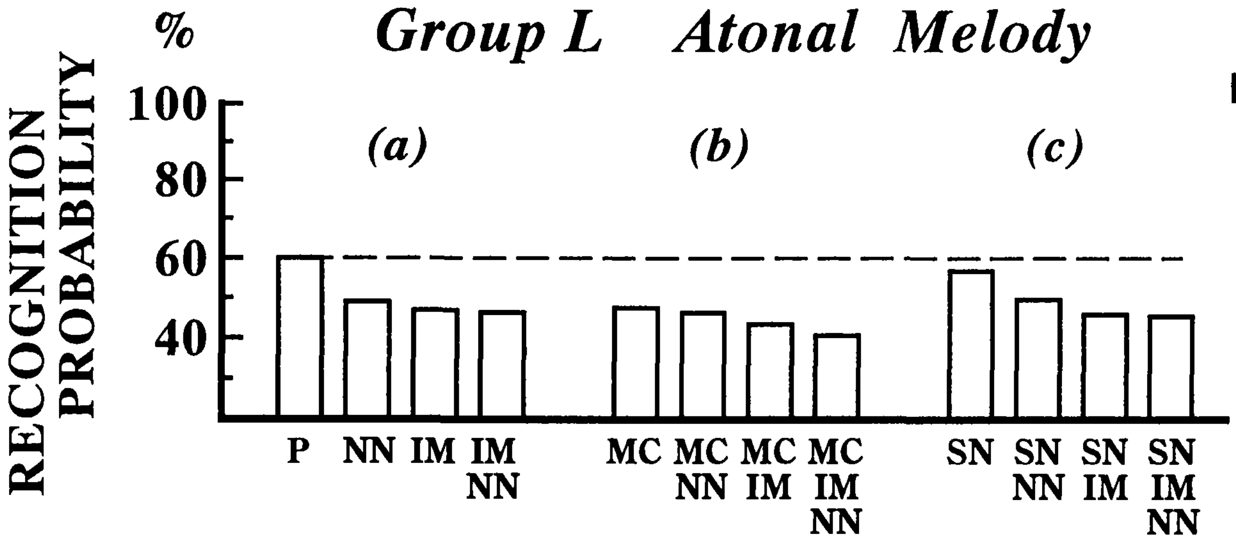


Fig.6-4

INTERFERENCE CONDITION

CHAPTER 5

EXPERIMENTAL TESTS OF MOTOR ENCODING STRATEGY FOR PITCH

EXPERIMENT 7

Finger Movement Strategy

Method

EXPERIMENT 7-1

EXPERIMENT 7-2

EXPERIMENT 7-3

Results and Discussion

EXPERIMENT 7

Finger Movement Strategy

In all of the previous Experiments 1 to 6 and in Experiments 8 and 9, subjects were instructed that they should not employ finger movements analogous to playing the piano to encode pitches. However, for the subjects who are highly trained in playing instruments, the development of performance skill on an instrument could be interpreted as leading to the development of a kind of auro–motor (auditory–motor) coordination that allows the musicians to reproduce immediately musical patterns that they experience as auditory images (Baily, 1985). That is, auditory imagery seems to be reinforced by the external representation of motor (kinesthetic) images. For the piano player, this intermodal coordination between auditory pitch information and finger movements has been experienced in relationship to an internal representation of the keyboard's spatial positions. Baily (1985) refers to this internal representations as a spatio–motor representation.

There is a series of studies on memory for action events (e.g., Cohen, 1981, 1983). The subject's task is to enact instructions (Subject–Performed Tasks : SPTs). The SPTs paradigm is designed to investigate the effect of encoding variables on the action events.

Some researchers suggest that motor enactment is related to visual image. Levin (1976) suggested that enactment facilitates memory because it leads subjects to form relevant visual images or visual mediation. Sasaki and Watanabe (1983) investigated the spontaneous writing "Kusho" behavior in Japanese. It was defined as writinglike finger movement without any physical and visible trace. The results showed that there are two types of this behavior ; (a) writing on any parts of body (e.g., knee, or palm etc.) with visual monitoring, and (b) writing in the air without visual monitoring. In this experiment, Kanji graphemes were presented in a visual or auditory form. In the visual presentation, two types of Kusho showed highly effect, however in the auditory condition, Kusho writing in the air without visual monitoring revealed no effect.

Other researchers suggest that motor enactment is independent of, and more dominant than, visual image. Saltz & Donnenwerth–Nolan (1981) suggested that many words and objects have motor implications, and that motor enactment can facilitate retention of sentences, and this facilitation is related to the storage of motor images or trace as distinct from visual images or verbal mediators (e.g., Engelkamp, 1990 ; Saltz & Dixon, 1982 ; Zimmer & Engelkamp, 1985, 1989).

The purpose of Experiment 7 was to investigate whether finger movements would be an effective strategy for encoding pitches of melodies (Mikumo, 1991a, 1992a,b,c, 1994b,c).

In a preliminary experiment, seven students highly trained in music (Group H) and seven less well musically trained subjects (Group L) were instructed to make recognition judgments of melodies following a 10–sec retention interval. Both the standard and comparison melodies were 6–tone series which had either a high–tonality structure (tonal) or a low–tonality structure (atonal). There were two encoding conditions : In Session 1, the subjects were not to employ the tapping strategy and were to use other strategies during the retention interval. In Session 2, they were to employ the tapping strategy. If the recognition performance was significantly higher with the tapping strategy than without, it could be concluded that tapping is an effective strategy for pitch encoding of melodies.

For each group recognition data (hit rate minus false–alarm rate [Woodworth & Schlosberg, 1954]) were analyzed in a two–way analysis of variance [2 Sessions X 2 Melody Types], with repeated measures on both factors. In Group H, there was a signifi–

cant main effect of Melody Type [$F(1,6)=39.70, p<.001$], and there was a significant interaction of Session X Melody Type [$F(1,6)=9.35, p<.05$]. In Group L, there was a main effect of Melody Type [$F(1,6)=6.66, p<.05$], and there was a marginal main effect of Session [$F(1,6)=2.97, p=.09$].

The results showed that recognition performance for highly musically trained subjects was a little higher when employing the tapping strategy with tonal melodies than when not employing tapping ; however, with atonal melodies, performance was disrupted by the tapping strategy. On the other hand, for less well musically trained subjects, performance was disrupted by the tapping strategy with both tonal and atonal melodies (Figure 7–1). The findings suggest that, for the highly trained subjects, employing the tapping strategy might be somewhat effective for pitch encoding of tonal melodies ; however in the other cases it appears to disturb pitch encoding. To encode pitch information, the tapping strategy might be too difficult and hence ineffectual for less well trained subjects, and for all subjects with atonal melodies.

The results obtained in the preliminary experiment presented several implications, which were investigated in Experiments 7–1 to 7–3. (a) Remembering 6–tone melodies is rather easy for highly trained subjects. In the preliminary experiment, their tonal recognition performance in the non–tapping session was relatively good. Therefore, in the case of 6–tone melodies, it would not be necessary to employ the laborious tapping strategy, and employing the tapping strategy would not necessarily be effective. Considering the above, the present experiments employed not only 6–tone but also 8– and 10–tone tonal melodies. (b) In the present experiments, finger movement was recorded by a computer, and it was possible to determine the rate and amount of tapping as well as accuracy. Moreover, to obtain detailed data of finger movement, ISI durations were varied. (c) In the present experiments, the latency between the end of the standard melody and the start of tapping could be measured and analyzed. Latency could well serve as an index of the important cognitive processes, because this period might involve various information processing stages. (d) In the preliminary experiment there was a blank interval during the ISI. To investigate the stability or robustness of the tapping strategy, an interfering melody or a series of musical note names were interpolated during the ISI, in the present experiments. (e) In the preliminary experiment some subjects were observed to employ the tapping strategy while muttering the pitches or the note names to themselves in a whisper. The present experiments were also designed to investigate whether a dual encoding strategy, such as pitch rehearsal or verbal encoding strategy along with the tapping strategy, is employed to encode pitches of melodies.

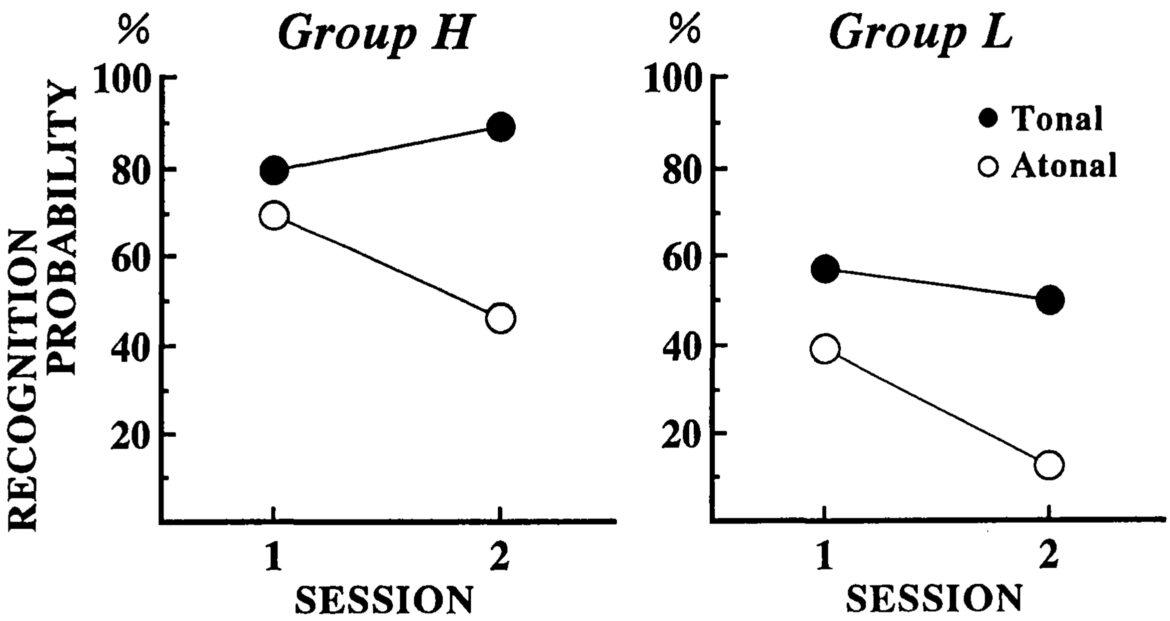


Figure 7-1.

Mean probability of correct recognition (hit rate minus false alarm rate) in Groups H and L for tonal and atonal melodies in Sessions 1 and 2 of preliminary experiment.

Notes.

Session 1 : not employing tapping, Session 2 : employing tapping

Experiments 7–1 to 7–3 employed subjects who were highly trained in music and used melodies which had strong tonal structure. To investigate the tapping strategy in more detail, the number of tones in a melody and the duration of the retention interval were varied, and finger movements were recorded by a computer. The retention interval involved a blank interval in Experiment 7–1, an interfering melody in Experiment 7–2, and a series of musical note names in Experiment 7–3. The methods for the three experiments are described first, followed by the results of all three.

Method

EXPERIMENT 7–1

Subjects

Twelve female subjects (average age, 21.3 years ; age range, 19–22 years) who were undergraduates majoring in music took part in this experiment. Each had had at least 15 years of formal training in playing the piano with an average of 17.5 years (range, 15–19 years). All of them were right hand dominant persons.

Materials

Each trial involved a standard series followed by a blank retention interval and then by a comparison series. Both the standard and the comparison series consisted of six, eight, or ten different tones taken from an equal–temperament scale, which ranged from B_3 (246.94 Hz) to E_5 (659.26 Hz). The duration of each tone was 700 msec (600 msec with 100–msec silence), so that the length of the 6–tone series was 4.2 sec, the 8–tone series was 5.6 sec, and the 10–tone series was 7.0 sec. All of the series were tonal melodies in a major key, and were high in tonal melodic structure according to conventional Western rules, and they involved a wide variety of contours.

For Experiments 7–1, 7–2 and 7–3, six lists were prepared, and each list involved 36 trials, so that a total of 216 trials were prepared. One list was used for one session. Each experiment had two sessions, and in each session there were 12 trials for each of the three types of melody lengths.

Three types of ISIs were prepared depending on the length of a melody. The short ISI (S), middle ISI (M), and long ISI (L) were, respectively, 1.5, 2.0, and 2.5 times as long as the melody length. Thus, in the 6–tone series, the short ISI was 6.3 sec, the middle ISI was 8.4 sec, and the long ISI was 10.5 sec. In the 8–tone series, the short ISI was 8.4 sec, the middle ISI was 11.2 sec, and the long ISI was 14.0 sec. In the 10–tone series, the short ISI was 10.5 sec, the middle ISI was 14.0 sec, and the long ISI was 17.5 sec.

Half of the comparison series were exactly the same as the standard series, and the other half were divided into two types that were different from the standard. "Contour–preserving" comparisons (C) were obtained by changing one of the pitches by two semitones (higher or lower), preserving the contour of the standard series, so that the exact pitch intervals were not preserved. "Exchanging" comparisons (E) were obtained by exchanging the order of two successive pitches of the standard series, so that both the contour and pitch intervals were somewhat different from those of the standard series, while the set of pitches was preserved. In both C and E comparisons, neither the first nor the last pitch of the standard series was altered, and the position of the change was counterbalanced among stimuli.

The tones were generated by a NEC PC–9801 26k sound synthesizer board installed in a NEC PC–9801 RX computer, and they were recorded on tape and played over high–quality sound reproduction equipment. All tones were adjusted to be equal in loudness (approxi–

mately 50 dB SPL), and were presented via a loudspeaker.

Procedure

Subjects were told that this was an experiment on memory for melodies. In each of the 36 trials, at first they would see a trial number on a computer display, then hear a first melody (standard series) which would be followed by a retention interval of varying length, and then they would hear a second melody (comparison series). The next trial would begin fifteen seconds after the second melody. The ISI length would change from trial to trial. The first and the second melodies would consist of six , eight, or ten tones. The subject's task would be to judge whether the two melodies are the same or different in pitch by answering orally, and rate their decision on a five–point confidence scale with responses of "very sure yes (or no)", "fairly sure yes (or no)", "unsure yes (or no)", "fairly unsure yes (or no)", "very unsure yes (or no)".

In the first session, the subjects were instructed to try to memorize or retain melodies without employing the tapping strategy, but rather by using other internal strategies during the retention interval. After the session, they were to report their introspections about what type of encoding strategies they used. In the second session, the subjects instructed to employ the tapping strategy, moving their fingers in a way analogous to playing the piano, during the retention interval. Considering the view of encoding specificity (Tulving & Thomson, 1973), they were to continue the finger movements while hearing the comparison series. The second session was performed thirty minutes after the first session. Three practice trials with feedback for each length of melody were given to the subjects prior to each session. If the session in which tapping is allowed is performed before the session in which tapping is not allowed, it would be difficult to eliminate the effect of motor learning in the latter session, because the motor sense of tapping tends to linger. This is why the two sessions were presented in fixed order to all the subjects.

As described above, six lists were prepared for Experiments 7–1, 7–2, and 7–3. These lists were counterbalanced between the subjects for each session in each experiment (Experiments 7–1, 7–2 and 7–3), so that the overall difficulty of the six sessions would be equal (Table 2). In each session, there were 12 six–tone, 12 eight–tone and 12 ten–tone trials. In three lengths of melody, 4 trials were run in each duration of ISI, in half of which (2 trials) the comparison series were exactly the same as the standard series, and in the other half (2 trials), the comparison series were different from the standard series (C, E). The three durations of ISI were mixed, and the order of presentation of all 36 trials in each session was randomized. Tapping movements of each finger of the right hand were recorded using a computer. The tip of each finger was fitted with a small piece of aluminum foil connected to the computer through fine leads. When each finger tapped on an aluminum board the number of times and the rate of touches were recorded.

EXPERIMENT 7–2

Subjects

The same subjects as those in Experiment 7–1 were employed.

Materials

Each trial involved a standard series followed by an interfering melody and then by a comparison series. An interfering melody consisted of three to nineteen tones taken from an equal–temperament scale, which ranged from B_3 (246.94 Hz) to E_5 (659.26 Hz), and the duration of each tone was 700 msec (600 msec with 100–msec silence).

The number of tones in the interfering melody depended on the duration of the ISI for each length of the melody. The intervals between the standard series and the first interfering tone, and between the last interfering tone and the comparison series, were each 2.1 sec. Thus, in the 6–tone series, therefore, there were 3 tones during the short ISI (6.3 sec), 6 tones during the middle ISI (8.4 sec), and 9 tones during the long ISI (10.5 sec). In the 8–tone series, there were 6 tones during the short ISI (8.4 sec), 10 tones during middle ISI (11.2 sec), and 14 tones during the long ISI (14.0 sec). In the 10–tone series, there were 9 tones during the short ISI (10.5 sec), 14 tones during the middle ISI (14.0 sec), and 19 tones during the long ISI (17.5 sec). Four interfering melodies were prepared for each case described above, so that a total of 36 interfering melodies were totally prepared, and all of them were tonal melodies in a major key.

Procedure

This experiment was performed two months after Experiment 7–1. The subjects were instructed that, in each of 36 trials, they would first see a trial number, then hear a first melody which would be followed by an interfering melody of varying length during the retention interval, and then they would hear a second melody. Thirty–six interfering melodies were employed in Sessions 1 and 2.

In the first session, the subjects were instructed that they were not to employ the tapping strategy, but rather use any other strategies, and that in the second session they were to employ the tapping strategy. The other aspects of the method were the same as in Experiment 7–1.

EXPERIMENT 7–3

Subjects

The same subjects as those in Experiments 7–1 and 7–2 were employed.

Materials

Each trial involved a standard series followed by a series of musical note names and then by a comparison series. The series of musical note names consisted of three to nineteen words chosen randomly from seven note names (Do, Re, Mi, Fa, Sol, La, Si). The series were delivered in a monotonous pitch in a male voice, and the duration of each word was about 700 msec (600 msec with 100-msec silence).

The number of words in a series of note names were the same as the number of tones in an interfering melody in Experiment 7-2, depending on the ISI duration for each length of the melody. In the 6-tone series, therefore, there were 3 words during the short ISI, 6 words during the middle ISI, and 9 words during the long ISI. In the 8-tone series, there were 6 words during the short ISI, 10 words during middle ISI, and 14 words during the long ISI. In the 10-tone series, there were 9 words during the short ISI, 14 words during the middle ISI, and 19 words during the long ISI. Four series of note names were prepared in each case described above, so that a total of 36 series was prepared.

Procedure

This experiment was performed two months after Experiment 7-2. The subjects were instructed that in each of 36 trials, they would first see a trial number, then hear a first melody which would be followed by a series of musical note names of variable length during the retention interval, and then they would hear a second melody. Thirty-six series of note names were employed in Sessions 1 and 2. Other aspects of the method were the same as in Experiments 7-1 and 7-2.

Table 2

Six lists for each Session in Experments 7−1, 7−2, & 7−3.

	Experiment 7−1		Experiment 7−2		Experiment 7−3	
Session	1	2	1	2	1	2
Subjects						
1, 2	List 1	List 2	List 3	List 4	List 5	List 6
3, 4	List 2	List 3	List 4	List 5	List 6	List 1
5, 6	List 3	List 4	List 5	List 6	List 1	List 2
7, 8	List 4	List 5	List 6	List 1	List 2	List 3
9, 10	List 5	List 6	List 1	List 2	List 3	List 4
11, 12	List 6	List 1	List 2	List 3	List 4	List 5

Experiment 7−1 : blank interval condition

Experiment 7−2 : interfering melody condition

Experiment 7−3 : interfering note names condition

Session 1 : not employing tapping

Session 2 : employing tapping

Results and Discussion

The design of these experiments had four factors : 2 types of Sessions (Session 1 : not employing tapping / Session 2 : employing tapping), 3 types of melody lengths (6 tones / 8 tones / 10 tones), 3 types of ISI durations (short / middle / long), and 2 types of comparison series (Contour–preserving / Exchanging). All of them employed the same subjects.

In these experiments the following response measures were calculated: the recognition probability for each melody length (Analyses 1 & 2), and for each ISI duration (Analysis 3) ; the discriminability index (d') (Analyses 1, 2, & 3) ; the number of incorrect response for each subject in each session (Analysis 2) ; the false–alarm rate for each comparison type (Analysis 4) ; and the amount and rate of tapping, and the latency between the end of the standard series and the start of finger movements (Analysis 5).

Analysis 1

Stability or robustness of the motor encoding strategy and the impact of the melody lengths in Experiment 7–1 and Experiment 7–2. Recognition probability data (hit rate minus false–alarm rate) were analyzed in a three–way analysis of variance [2 Experiments X 2 Sessions X 3 Melody Lengths], with repeated measures on all three factors. There were significant main effects of Experiment, Session, and Melody Length [$F(1,11)=12.57, p<.005$; $F(1,11)=45.90, p<.001$; $F(2,22)= 80.51, p<.001$]. There was also a marginal interaction of Experiment X Session [$F(1,11)=3.30, p<.10$]. These results are shown in Figure 7–2 and Table 3.

The main effect of Experiment indicates that the performance in Experiment 7–1 was superior to that in Experiment 7–2. The main effect of Session indicates that in both Experiments 7–1 and 7–2, subjects performed better when they employed the tapping strategy than when they did not. The main effect of Melody Length indicates that the subjects performed gradually worse as Melody Length increased. The marginal interaction of Experiment X Session indicates the following : When several tones were interpolated during the ISI in Experiment 7–2, and subjects did not employ the tapping strategy in Session 1, then their performance was significantly worse than it was in Session 1 of Experiment 7–1 (especially in the case of the 8–tone and the 10–tone series). However when they employed the tapping strategy in Experiment 7–2, their performance was not worse than in Session 2 of Experiment 7–1.

Subjects reported on their internal encoding strategies in Session 1 of Experiment 7–1 as follows : verbal encoding strategy with the names of the musical notes (6 subjects) ; pitch rehearsal or humming strategy (2 subjects) ; and visualizing strategies, in which the tones were visualized in their image, as a melodic contour (1 subject), on a staff notation (2 subjects), or on a keyboard (1 subject). For each subject, these encoding strategy would be dominant or individual in ordinary condition, because these internal encoding strategies were obtained without any interfering stimuli during the ISI.

Therefore, these findings suggest that, for subjects highly trained in music, employing the external motor strategy of tapping could be more effective than using other internal strategies to encode and retain the pitches of the melodies, especially within a situation in which there were interfering melodies. Moreover, as shown in Figure 7–2 and Table 3, the effect of tapping represented by the difference between the data of Session 1 and Session 2 increased somewhat as melody length increased in both of the experiments.

Discriminability indices (***d'***) were calculated for three types of melody lengths in the two sessions of the two experiments (Table 3), by using the ten–category confidence scale in which "Same" and "Different" judgments each had five levels of response. It was evident from the *d'* analysis that discriminative power in Same–Different decisions of Session 2 was consistently better than that of Session 1 in both experiments (except for the 6–tone series in Experiment 7–1). The differences of the *d′* data between the sessions were tested by a Z–test as suggested by Gourevitch and Galanter (1967). There were significant differences between the sessions in the 8–tone and in the 10–tone series of Experiment 7–1, and in the 6–tone series of Experiment 7–2 ($p<.001$) ; and between the sessions in the 8–tone and in the 10–tone series of Experiment 7–2 ($p<.01$). These results support the findings of the recognition probability data described above.

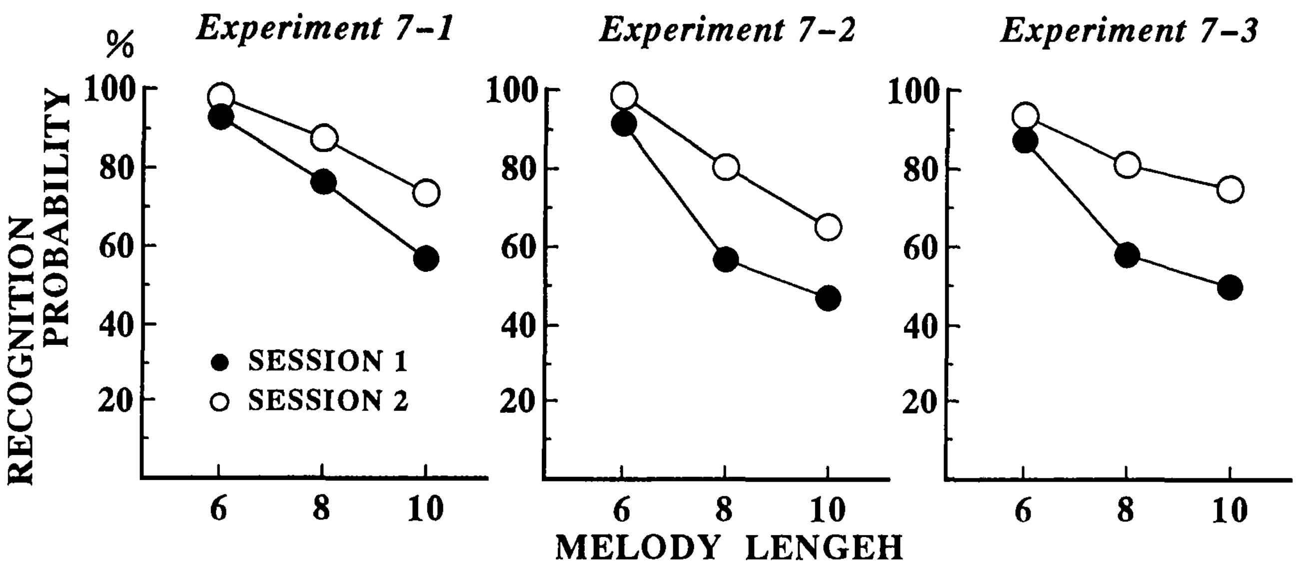


Figure 7–2.

Mean probability of correct recognition for each melody length in two sessions of Experiments 7–1, 7–2, & 7–3 (N=8). The tapping effect is represented by the distance between the white circle (Session 2) and the black circle (Session 1).

Notes.

Experiment 7–1 : blank interval condition, Experiment 7–2 : interfering melody condition, Experiment 7–3 : interfering note names condition, Session 1 : not employing tapping, Session 2 : employing tapping.

Table 3

Recognition probability and d' for each ISI duration for each melody length of the two sessions in Experiments 7−1, 7−2, & 7−3.

Melody Length	Experiment 7−1						Experiment 7−2						Experiment 7−3 (N=8)					
	Session 1			Session 2			Session 1			Session 2			Session 1			Session 2		
ISI	6−tone	8−tone	10−tone	6−tone	8−tone	10−tone	6−tone	8−tone	10−tone	6−tone	8−tone	10−tone	6−tone	8−tone	10−tone	6−tone	8−tone	10−tone
S (ML × 1.5)	5.3 [b]	3.5	1.7	6.9	3.9	1.7	5.2	1.8	1.9	6.9	1.6	1.8	3.9	2.5	0.9	5.4	2.7	1.9
M (ML × 2.0) [a]	5.6	3.1	1.8	6.9	4.3	4.0	2.8	1.6	0.8	6.1	1.9	3.0	2.2	1.8	1.4	6.9	2.7	2.8
L (ML × 2.5)	6.9	1.9	1.5	6.9	4.4	5.3	4.6	0.8	0.7	6.9	2.8	3.2	5.5	2.3	0.7	5.9	3.2	2.5
S + M + L	5.7	2.5	1.7	6.9	4.2	3.4	4.5	1.3	1.3	6.6	2.5	2.4	4.9	2.3	1.1	5.9	2.6	2.2
Probability (%) [c]	93.0	76.3	57.0	98.0	87.5	73.6	91.7	57.0	47.2	98.6	80.5	65.3	87.5	58.3	50.0	93.8	81.3	75.0

a) M (ML × 2.0) means that the middle ISI is twice as long as the melody length.

b) Discriminative power in the Same−Different decision for 6−tone melodies is so strong that the d' can not be calculated. Therefore d's for 6−tone melodies are fixed using the method suggested by Macmillan & Creelman (1991).

c) Recognition probability data were calculated by the method of hit rate minus false alarm rate.

Experiment 7−1 : blank interval condition, Experiment 7−2 : interfering melody condition, Experiment 7−3 : interfering note names condition
Session 1 : not employing tapping, Session 2 : employing tapping

Analysis 2

The number of incorrect responses and the impact of melody lengths in Experiment 7–3 compared with Experiment 7–1 and Experiment 7–2. The number of incorrect responses in 36 trials of each session for each subject is shown in Table 4. Table 4 shows that, for almost all the subjects the number of incorrect responses of Session 2 was consistently lower than that of Session 1 in Experiments 7–1 and 7–2. In other words, in Experiments 7–1 and 7–2, tapping strategy was effective for encoding pitches of melodies, even when an interfering melody was interpolated during the ISI. In Session 2 of Experiment 7–2, after hearing the standard series, if subjects have employed tapping strategy along with muttering or singing pitches to themselves in a whisper or rehearsing them mentally, interfering melodies would disrupt tapping with pitch rehearsal and recognition performance. But the results showed that the number of incorrect responses was consistently lower in Session 2 than in Session 1 in Experiment 7–2, it is considered that the tapping strategy could be employed without muttering or singing pitches ; that is, tapping strategy could be employed independently of pitch rehearsal strategy for all twelve subjects.

On the other hand, in Experiment 7–3, for eight subjects, the number of incorrect responses was lower or about the same in Session 2 as in Session 1, whereas for the remaining four subjects, errors were considerably higher in Session 2 than in Session 1. For these four subjects, the tapping strategy was effective in Experiments 7–1 and 7–2, while in Experiment 7–3, the series of note names disrupted tapping and recognition performance in Session 2. It is considered that it likely that those four subjects employed the tapping strategy along with muttering the note names to themselves in a whisper or rehearsing them mentally. That is, they may have employed a dual–coding strategy, in that when they employed the tapping strategy at the same time, they continued to use verbal encoding strategy.

Table 4

The number of the incorrect responses for each subject in each session (36 trials) of Experiments 7-1, 7-2, & 7-3.

Subjects	Experiment 7-1		Experiment 7-2		Experiment 7-3	
	Session 1	Session 2	Session 1	Session 2	Session 1	Session 2
5	2	0	7	4	5	5
11	2	1	8	2	1	2
3	2	4	6	2	11	3
6	6	5	9	4	11	4
2	5	4	5	4	4	3
4	8	1	7	5	6	4
7	5	2	6	1	5	0
10	5	1	7	1	7	3
1	4	2	7	4	5	12
8	6	5	7	4	5	8
9	6	1	6	4	2	9
12	2	2	7	4	1	6

Experiment 7-1 : blank interval condition
Experiment 7-2 : interfering melody condition
Experiment 7-3 : interfering note names condition
Session 1 : not employing tapping
Session 2 : employing tapping

Comparing the numbers of incorrect responses for each subject for Session 1 of the three experiments, the results are consistent with subjects' introspections concerning internal encoding strategies in Session 1 of Experiment 7-1. In the case of two Subjects 5 and 11, incorrect responses in Session 1 of Experiment 7-2 were especially higher ; that is, the disruptive effect of the interfering melodies was large. This suggests that their dominant encoding strategy was pitch rehearsal as they reported in Experiment 7-1. In the case of two other Subjects 3 and 6, incorrect responses in Session 1 of Experiment 7-3 were especially higher ; that is, the disruptive effect of the series of note names was large. This suggests that they were using the verbal encoding strategy with note names as they reported in Experiment 7-1. In the case of Subjects 2, 4, 7 and 10, there was little difference among their incorrect responses in Session 1 of the three experiments ; that is, there was little disruptive effect of the interfering melodies or the series of note names. This suggests that these subjects were using an encoding strategy other than pitch rehearsal or verbal encoding. They reported that they used a visualizing strategy in which tones were visualized in their image, as a melodic contour (Subject 2), on a staff notation (Subjects 4 and 10), or on a keyboard (Subject 7). After Session 2 of Experiment 7-1, however, Subjects 7 and 10 reported that if they had been permitted to employ any kind of strategy in Session 1, their dominant strategy for pitch encoding might have been the tapping strategy. In the case of Subjects 1, 8, 9, and 12, they reported in Experiment 7-1 that they also used a verbal encoding strategy with note names. When they employed the tapping strategy to encode pitches, they could not employ it independently of the verbal encoding strategy ; they may have employed a dual-coding strategy as described above. Furthermore, they may have rehearsed note names along with their pitches, therefore, it is possible that they employed triple-encoding strategy ; tapping strategy accompanied by note names with pitches, as described in Experiment 6. As reported in Experiment 7-1, although there were individual differences for encoding strategies, especially for six Subjects 5, 11, 3, 6, 2, and 4, more direct evidence for the effectiveness of finger movements came from the instruction to employ motor encoding strategy.

The recognition probability data of eight subjects in Experiment 7–3 were analyzed in a two–way analysis of variance [2 Sessions X 3 Melody Lengths], with repeated measures on both factors. In this analysis the data of the four subjects who employed a multi–coding strategy were excluded. The recognition probability data of these eight subjects in Experiment 7–3 are also shown in Figure 7–2 and Table 3. There were significant main effects of Session and Melody Length [$F(1,7)=7.84, p<.05$; $F(2,14)=14.99, p<.001$].

The main effects of Session and Melody Length were consistent with the findings of Analysis 1. When a series of note names were interpolated during the ISI, the tapping strategy would be also effective for encoding pitch information. As shown in Figure 7–2 and Table 3, the tapping effect, which is represented by the difference between the data of Session 1 and Session 2, becomes gradually stronger as melody length increased. Discriminability indices (*d'*) for Experiment 7–3 are also shown in Table 3. Both recognition performance and *d'* for Experiment 7–3 with the eight subjects were similar to the results of Experiment 7–2.

The recognition probability data of the four subjects who employed a multi–coding strategy were analyzed in a three–way analysis of variance [3 Experiments X 2 Sessions X 3 Melody Lengths], with repeated measures on all three factors. There was a significant main effect of Melody Length [$F(2,6)=45.07, p<.001$], and there was a significant interaction of Experiment X Session [$F(2,6)=21.26, p<.005$]. There was a marginal main effect of Experiment [$F(2,6)=3.82, p<.09$]. These results were shown in Figure 7–3.

The significant interaction of Experiment X Session indicates that in Experiments 7–1 and 7–2 recognition performance in Session 2 was superior to that in Session 1 (Experi–

ment 7–1, $p<.05$; Experiment 7–2, $p<.01$), while in Experiment 7–3, in contrast, Session 1 was superior to Session 2 ($p<.01$). The marginal main effect of Experiment indicates that recognition performance in Experiment 7–1 was superior to that in Experiment 7–2 ($p<.05$), and there was no significant difference between Experiment 7–2 and Experiment 7–3. These findings support the interpretation of Table 4.

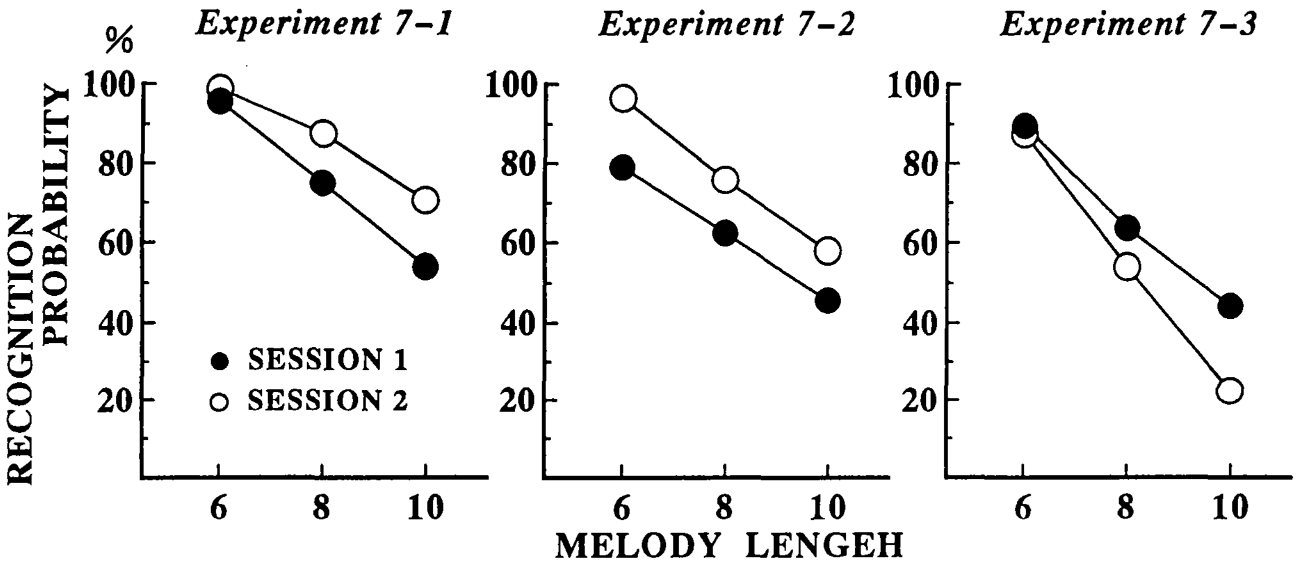


Figure 7–3.

Mean probability of correct recognition for the four subjects, for each melody length in two sessions of Experiments 7–1, 7–2, & 7–3. The tapping effect is represented by the distance between the white circle (Session 2) and the black circle (Session 1).

Notes.

Experiment 7–1 : blank interval condition, Experiment 7–2 : interfering melody condition, Experiment 7–3 : interfering note names condition, Session 1 : not employing tapping, Session 2 : employing tapping.

Analysis 3

The impact of ISI durations. Recognition probability data for Experiments 7–1 and 7–2 were analyzed in a three–way analysis of variance [2 Experiments X 2 Sessions X 3 ISIs], with repeated measures on all three factors. As in Analysis 1 there were significant main effects of Experiment and Session [$F(1,11)=11.60$, $p<.01$; $F(1,11)=50.85$, $p<.001$], and a marginal interaction of Experiment X Session [$F(1,11)=3.28$, $p<.10$]. In addition, there was a tendency toward an interaction of Session X ISI [$F(2,22)=2.08$, $p=0.14$]. This tendency indicates that the recognition probability in Session 1 became gradually worse as ISI increased, whereas in Session 2, it became gradually better. From a two–way analysis of variance [2 Sessions X 3 ISIs] for each melody length in each experiment, this tendency (an interaction of Session X ISI) was especially noticeable in the case of the 10–tone series, as shown in Figure 7–4 (10–tone series : Experiment 7–1, $F(2,22)=3.36$, $p<.06$; Experiment 7–2, $F(2,22)=3.12$, $p<.07$]. In both Experiments 7–1 and 7–2 the significance levels between long and short ISIs in Session 2 were $p<.05$, whereas in Session 1 of both experiments there were no significant differences between short and long ISI (by Newman–Keuls Method).

These findings suggest that when subjects used internal encoding strategies their memory of the standard series became gradually obscure as ISI duration increased. However, when subjects employed the external motor encoding strategy, their memory became gradually elaborated as ISI duration increased.

Discriminability indices (***d'***) were calculated for the three ISI durations for each melody length in the two sessions of Experiments 7–1 and 7–2 (Table 3). It was evident from the *d'* that discrimination became worse as ISI increased in Session 1, while it became better as ISI increased in Session 2 in both experiments, except for the 6–tone series. The differences in d' among the three ISI durations for the 10–tone series were tested by a Z–test. In Session 2 of Experiment 7–1, there were significant differences between the middle ISI and the short ISI ($p<.001$), and between the long ISI and the middle ISI ($p<.01$) ; in Session 1, there were no differences among any of the three ISI durations. In Session 2 of Experiment 7–2, there was a significant difference between the middle ISI and the short ISI ($p<.02$), and there was no significant difference between the long ISI and the middle ISI ; in Session 1, there was a significant difference between the short ISI and the middle ISI ($p<.05$), and there was no significant difference between the middle ISI and the long ISI. These results support the findings of the recognition probability for the ISI duration described above.

In Experiment 7–3, the recognition probability data were analyzed in two two–way analyses of variance for the two subject groups [2 Sessions X 3 ISIs], with repeated measures on both factors. Both for the eight subjects discussed above and the four subjects who employed a multi–coding strategy, there were main effects of Session. Recognition performance for the eight subjects in Session 2 was higher than in Session 1 [$F(1,7)=7.84$, $p<.05$], while the reverse was true for the four subjects [$F(1,2)= 24.53$, $p<.001$]. However, there were no interactions between Session X ISI. Discriminability indices (d') for Experiment 7–3 with the eight subjects are also shown in Table 3.

These findings in Analyses 1, 2 and 3 suggest that when the retention interval was filled with an interfering melody or a series of note names, or as melody length or ISI duration increased, subjects would make an effort to utilize tapping in order not to forget the standard series, and rehearsing tapping could elaborate the encoding of pitches.

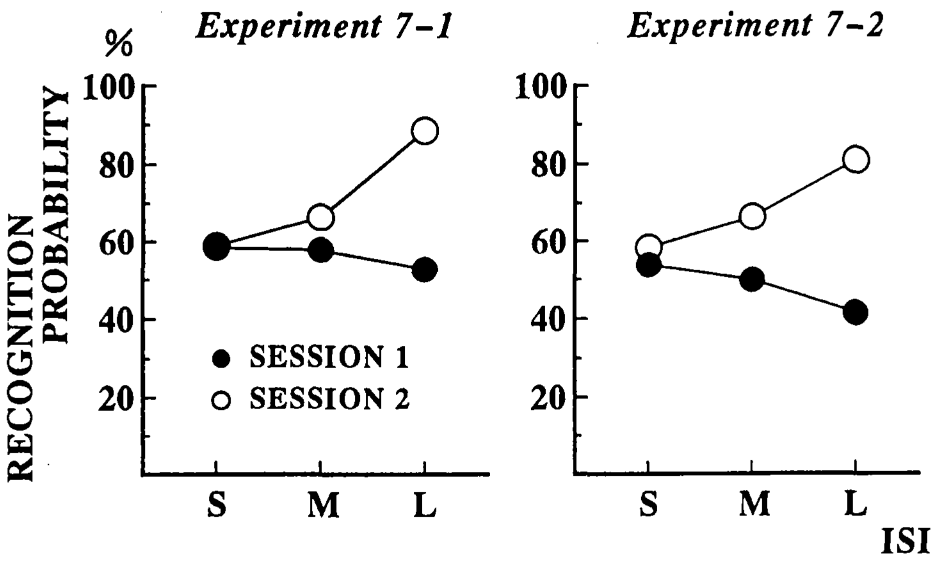


Figure 7-4.

Mean probability of correct recognition for each ISI in two sessions of Experiments 7-1 & 7-2 (10-tone series). The tapping effect is represented by the distance between the white circle (Session 2) and the black circle (Session 1).

Notes.

Experiment 7-1 : blank interval condition,
Experiment 7-2 : interfering melody condition,
Session 1 : not employing tapping, Session 2 : employing tapping.

Analysis 4

False-alarm rates for different Comparison Types. First, false- alarm rate data for Experiments 7-1 and 7-2 were analyzed in a three-way analysis of variance, [2 Experiments X 2 Sessions X 2 Comparison Types], with repeated measures on all three factors. There were significant main effects of Experiment, Session and Comparison Type [

$F(1,11)=67.53, p<.001$; $F(1,11)=60.95, p<.001$; $F(1,11)=41.46, p<.001$]. There were significant interactions of Experiment X Session and Experiment X Comparison Type [$F(1,11)=35.14, p<.001$; $F(1,11)=5.77, p<.05$].

The main effect of Experiment indicates that the false–alarm rate in Experiment 7–2 was higher than that in Experiment 7–1. The main effect of Session indicates that the false–alarm rate in Session 1 was higher than that in Session 2. The interaction of Experiment X Session indicates that in Session 1 the false–alarm rate in Experiment 7–2 was significantly higher than that in Experiment 7–1 ($p<.01$), whereas in Session 2 there was no significant difference between Experiments 7–2 and 7–1. Therefore, the main effects of Experiment and Session, and the interaction of Experiment X Session were consistent with the results of Analyses 1 and 3, described above.

The main effect of Comparison Type indicates that the false–alarm rate for C lures was higher than that for E lures. The interaction of Experiment X Comparison Type indicates that, for Cs the false–alarm rate in Experiment 7–2 was significantly higher than that in Experiment 7–1 ($p<.01$), and for Es there was also a significant difference between Experiment 7–2 and Experiment 7–1 ($p<.05$) ; but that the difference for Cs was larger than that for Es.

Second, false–alarm rate data were analyzed in three three–way analyses of variance for the three experiments: [2 Sessions X 3 Melody Lengths X 2 Comparison Types], with repeated measures on all three factors. In Experiment 7–1, there were significant main effects of Session, Melody Length, and Comparison Type [$F(1,11)=6.49, p<.05$; $F(2,22)=7.69, p<.01$; $F(1,11)=11.99, p<.01$]. There was a significant interaction of Melody Length X Comparison Type [$F(2,22)=4.68, p<.05$], and there was a marginal interaction of Session X Melody Length [$F(2,22)=2.50, p<.10$]. In Experiment 7–2, there were significant main effects of Session, Melody Length, and Comparison Type [

$F(1,11)=57.56, p<.001$; $F(2,22)=56.57, p<.001$; $F(1,11)=28.64, p<.001$], and there were significant interactions of Session X Melody Length, and Melody Length X Comparison Type [$F(2,22)=17.87, p<.001$; $F(2,22)=5.92, p<.01$]. In Experiment 7–3 for the eight subjects discussed above, there were significant main effects of Melody Length and Comparison Type [$F(2,14)=5.30, p<.05$; $F(1,7)=7.09, p<.05$], and there was a significant interaction of Melody Length X Comparison Type [$F(2,14)=3.74, p<.05$].

The main effects of Session indicate that the false–alarm rate in Session 1 was higher than that in Session 2. The main effects of Melody Length indicate that the false–alarm rate became gradually higher as Melody Length increased (significance levels between 10– and 6–tone series were $p<.05$ in Experiment 7–1, $p<.01$ in Experiment 7–2, and $p<.05$ in Experiment 7–3). The interactions of Session X Melody Length indicate that the false-alarm rate in Session 1 became gradually higher as Melody Length increased (significance levels between 10– and 6–tone series were $p<.05$ in Experiment 7–1, and $p<.01$ in Experiment 7–2), whereas in Session 2, there were no significant differences as Melody Length increased. That is, the tapping effect represented by the difference between the data of Sessions 1 and 2, became stronger as melody length increased. Therefore the main effects of Session and Melody Length, and the interaction of Session X Melody Length support the findings of Analyses 1 and 3. The main effects of Comparison Type were consistent with the results in the first analysis under Analysis 4, described above.

The interactions of Melody Length X Comparison Type indicate the following : for Cs false–alarm rate became gradually higher as Melody Length increased (significance levels between 10– and 6–tone series were $p<.05$ in Experiment 7–1, $p<.01$ in Experiment 7–2, and $p<.01$ in Experiment 7–3) ; whereas for Es there were no significant differences attributable to Melody Length in Experiment 7–1, and relatively little differences in Experiments 7–2 and 7–3 (significance levels between 10– and 6–tone series were $p<.05$ in both Experiments 7–2 and 7–3).

Third, false–alarm rate data were analyzed in three three–way analyses of variance for the three experiments : [2 Sessions X 3 ISIs X 2 Comparison Types], with repeated measures on all three factors. In Experiment 7–1, there were significant main effects of Session and Comparison Type [$F(1,11)=15.40, p<.01$; $F(1,11)=9.16. p<.05$], and there were no significant interactions between any of the three factors. In Experiment 7–2, there were significant main effects of Session and Comparison Type [$F(1,11)=64.10, p<.001$; $F(1,11)=59.40, p<.001$]. There was a significant interaction of Session X Comparison Type [$F(1,11)=5.60, p<.05$], and there was a marginal interaction among all three factors [$F(2,22)=3.36, p<.06$]. In Experiment 7–3 for the eight subjects discussed above there was a significant main effect of Comparison Type [$F(1,7)=8.00, p<.05$], and there was a marginal interaction among all three factors [$F(2,14)=2.86, p<.09$].

The main effects of Session and Comparison Type were consistent with the results in the first and the second analyses under Analysis 4. The significant interaction of Session X Comparison Type in Experiment 7–2 indicates that false–alarm rate for Cs were significantly higher than for Es in Session 1 ($p<.01$), while in Session 2 there was no significant difference. The marginal interactions among all three factors in Experiments 7–2 and 7–3 indicate the following : In Session 1, false–alarm rate for Cs gradually decreased as ISI increased (significance levels between short and long ISIs were $p<.05$ in Experiments 7–2 and 7–3), while for Es, it increased gradually as ISI increased (significance levels between long and short ISIs were $p<.01$ in Experiment 7–2, and $p<.05$ in Experiment 7–3). In Session 2, however, there were no significant differences among the three ISIs for both Cs and Es (Figure 7–5). This tendency toward a marginal interaction appeared slightly in Experiment 7–1.

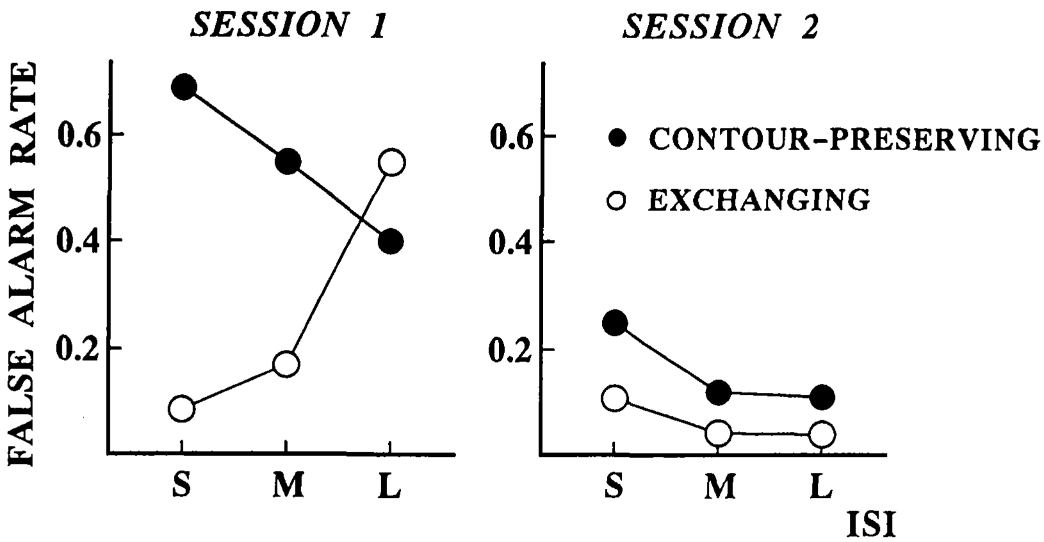


Figure 7-5.

Mean false-alarm rate for each ISI of two types of comparison series in Sessions 1 & 2 of Experiment 7-2.

Notes.

Experiment 7-2 : interfering melody condition,
Session 1 : not employing tapping, Session 2 : employing tapping.

As described above, C comparisons preserved the contour of the standard series, but the pitch set was not preserved. E comparisons preserved the pitch set of the standard, but contour was not preserved. If subjects were to use contour as a cue for recognition, they would tend to make false alarms to C comparisons which have the same contour as the standard. If subjects were to use the pitch set as a cue, they would tend to make false alarms to E comparisons which have the same pitch set as the standard.

In the first to third analyses for Comparison Type under Analysis 4, the following results were found : (a) The false–alarm rate for Cs was considerably higher in Experiments 7–2 and 7–3 than in Experiment 7–1. Therefore subjects appear to be using contour as a cue, especially when an interfering melody or a series of note names was interpolated during ISI. (b) The false–alarm rate for Cs increased gradually as melody length increased. Therefore subjects appear to depend more and more on contour as a cue, as melody length increased. Edworthy (1985) showed that performance was better in the contour than in the pitch interval task for melodies of up to nine tones in length, and there were no significant differences between tasks for the 11– and 13–tone melodies, but for the 15–tone melodies, performance was better in the pitch interval than in the contour task. Experiment 7 employed melodies of up to ten tones, therefore the result concerning contour was nearly consistent with Edworthy's findings. (c) In Session 1 the false–alarm rate for Es became gradually higher, but for Cs became gradually lower, as ISI duration increased ; while in Session 2 there were no significant effects of ISI. In other words, when not employing the tapping strategy subjects would use contour as a cue with short ISIs, whereas they would use pitch as a cue with long ISIs (e.g., Dowling & Bartlett, 1981). However, when subjects employed tapping strategy, the finger movement served as an effective means of encoding, even as ISI duration increased.

Analysis 5

Finger movements. The movement of each finger in tapping was recorded by a computer. Three response measures were calculated : (a) The number of times that the tapped pattern was repeated during the retention interval (ISI). (b) The rate of tapping for each finger. (c) The latency between the end of the standard series and the start of finger movement.

As shown in Figure 7–6, the number of times that the tapped pattern was repeated during the retention interval filled with interfering melodies (Experiment 7–2) or series of note names (Experiment 7–3) was more than that occurred in the blank retention interval (Experiment 7–1). In other words, as shown in Figure 7–7, the rate of tapping for each finger was faster in Experiments 7–2 and 7–3 than in Experiment 7–1. Moreover, the rate became faster as melody length increased and as ISI duration increased. For example, though it takes 700 msec per one tone of the stimulus melody, it took about 330 msec for each finger, corresponding to one tone of the stimulus melody, in the short ISI of the 6–tone melody in Experiment 7–1 ; that is, the rate of tapping was about twice the rate of the stimulus melody. On the other hand, it took about 180 msec for each finger in the long ISI of the 10–tone melody in Experiment 7–3 ; that is , the rate of tapping was about four times the rate of the stimulus melody. The data in Figures 7–5 and 7–6 support the interpretation of the Analyses 1, 2 and 3.

Sternberg, Monsell, Knoll, and Wright (1980) and Sternberg, Wright, Knoll, and Monsell (1980) reported a series of studies of the timing of rapid prespecified action sequences in speech and typewriting. Results showed that the mean latency for rapid reciting of a prespecified utterance increased approximately linearly with list length ($n<6$) ; the slope is 7.3 $\pm$ 0.7 msec/word (284.9 + 7.3(n–1) msec). The mean rate increased approximately linearly with list length ($n<6$) ; the slope is 180.4 + 12.0(n–1) msec/word for two–syllable words. The mean latency for typewriting of lists of letters by fingers on one hand increased approximately linearly with list length ($2<=n<=5$) ; the slope is 4.1 $\pm$ 1.3 msec/letter (226 + 4.1(n–1) msec). The mean rate increased approximately linearly with list length ($2<=n<=5$) ; the slope is 142.9+14.1(n–1) msec/letter with one hand. Therefore, Sternberg, et al. concluded that increments in sequence length increased the time to initiate the first response element (latency effect) and also increased the average time interval from one unit to the next (rate effect).

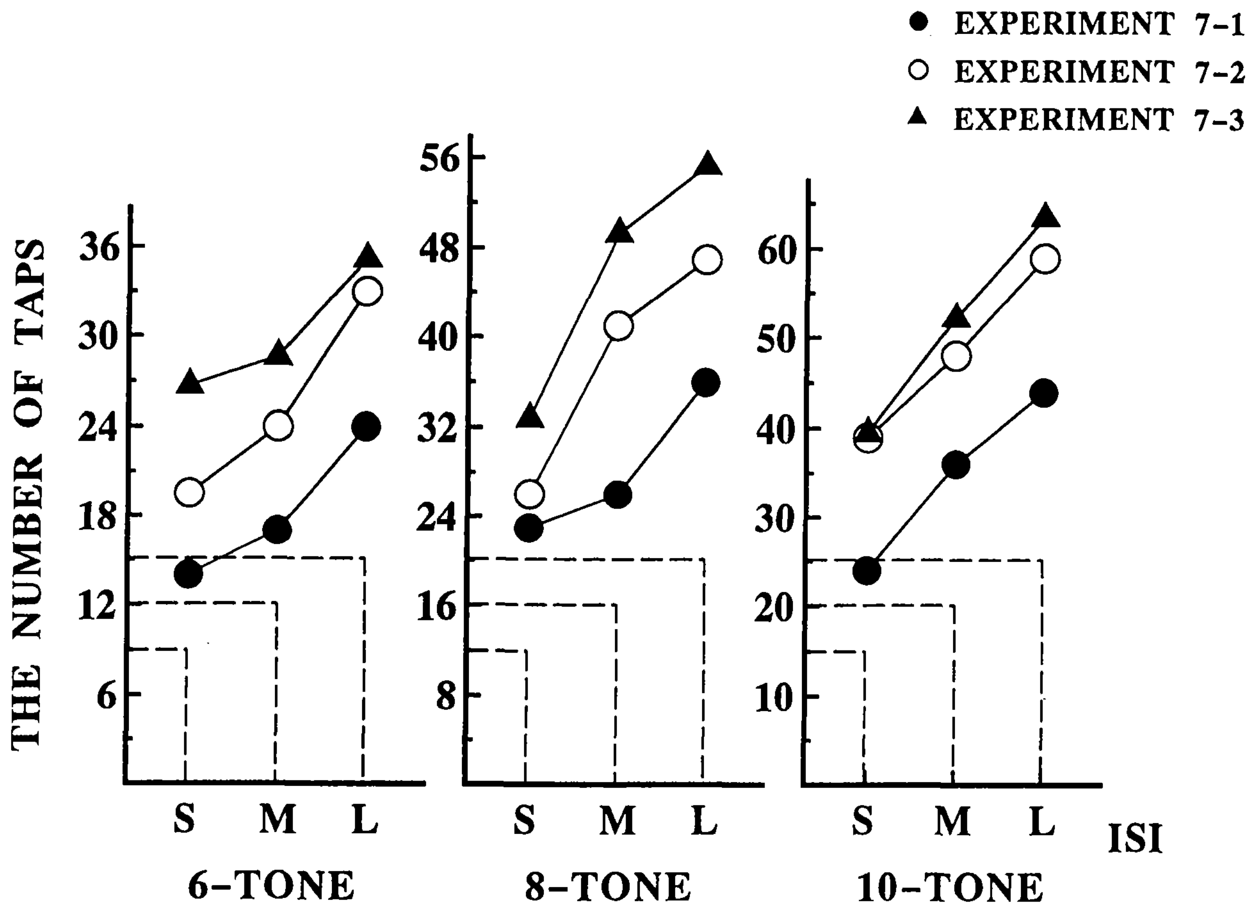


Figure 7-6.

The number of taps rehearsed during each ISI for each melody length in Experiments 7-1, 7-2, & Experiment 7-3 (N=8). Dotted line indicates that, for example, the middle ISI is twice as long as the standard series, therefore, if subjects tapped at the same rate as each tone of the standard series, they tapped twelve in the case of the 6-tone series. Twelve taps means that the 6-tone standard series was rehearsed twice; or 21 taps means that it was rehearsed three times plus 3 taps.

Notes.

Experiment 7-1 : blank interval condition, Experiment 7-2 : interfering melody condition, Experiment 7-3 : interfering note names condition.

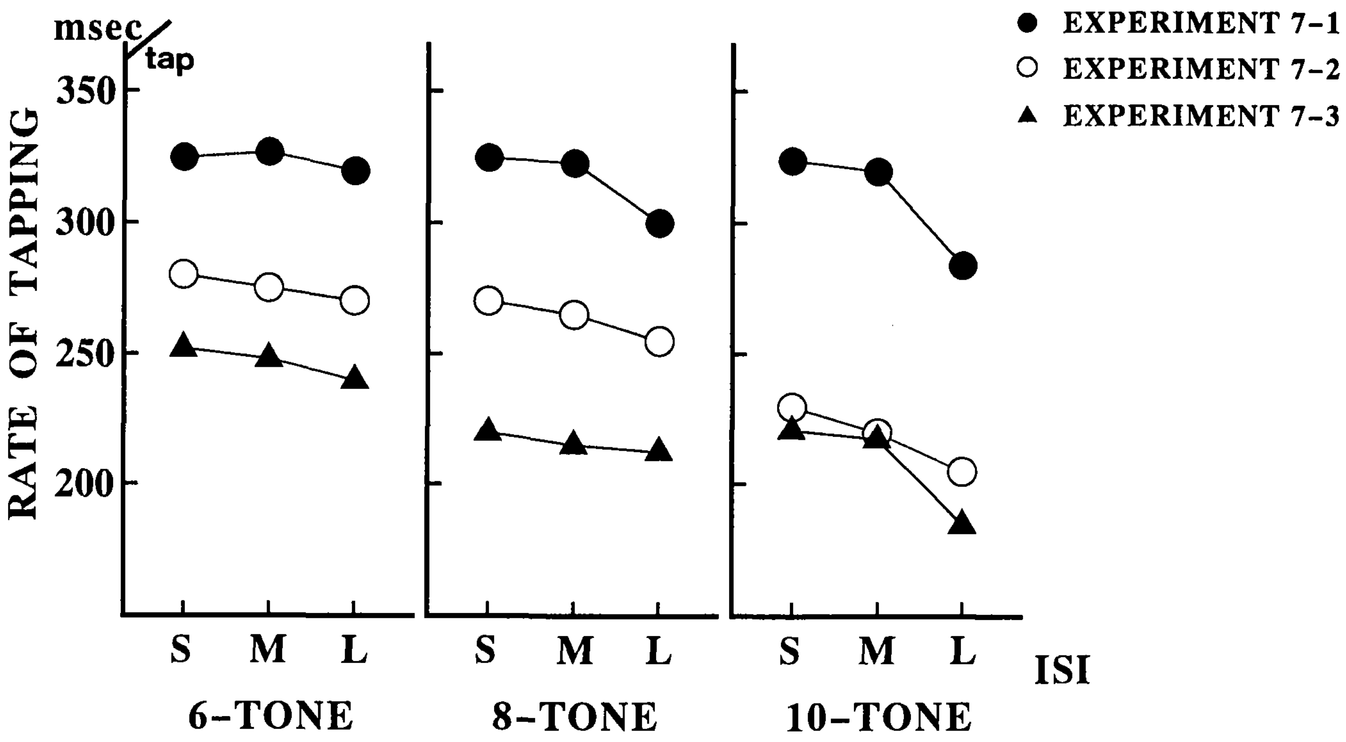


Figure 7-7.

The rate of tapping for each finger, for each melody length and each ISI in Experiments 7-1, 7-2, & Experiment 7-3 (N=8).

Notes. Experiment 7-1 : blank interval condition, Experiment 7-2 : interfering melody condition, Experiment 7-3 : interfering note names condition.

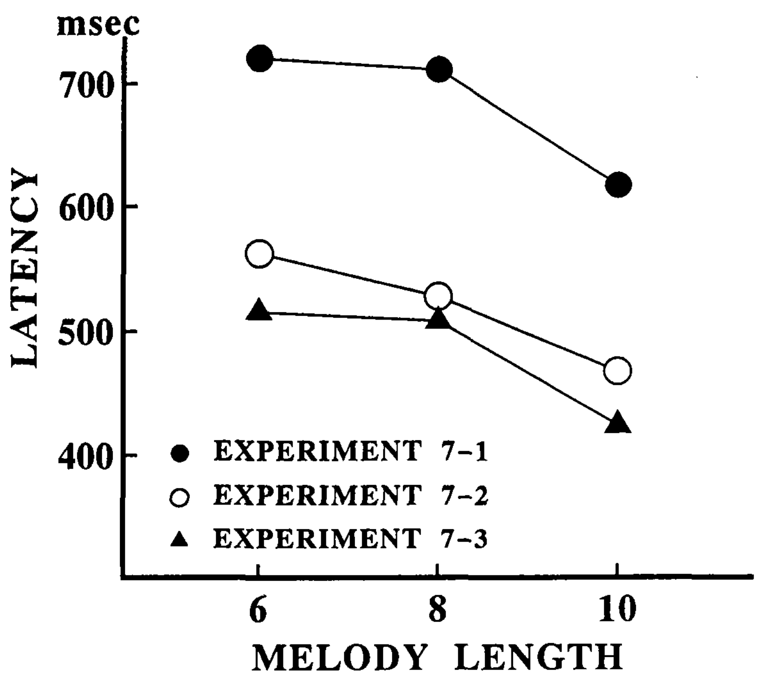


Figure 7-8.

Latency period between the end of the standard series and the start of tapping for each melody length in Experiments 7-1, 7-2, & Experiment 7-3 (N=8).

In contrast, in the present experiments, the rate of tapping for each finger became faster (Figure 7–7) and the latency became shorter (Figure 7–8) as melody length increased. Furthermore the latency in Experiments 7–2 and 7–3 was shorter than that in Experiment 7–1 ; tapping started earlier in Experiments 7–2 and 7–3 (interference conditions) than in Experiment 7–1 (blank–interval condition). This may indicate that in the challenging conditions in terms of melody length and interfering stimuli, the subjects may have been mentally prepared to start tapping earlier and to repeat the pattern as many times as possible, before the memory of the standard melody was disturbed by interfering stimuli. When the standard melody had ten tones, the latency was remarkably shorter. As shown in Figure 7–8, in the interference conditions subjects started tapping about 450 msec after they heard the last tone of the 10–tone standard melody. On the other hand, in the blank–interval condition, latency was about 610 msec for the 10–tone melody. Therefore the latency could be reduced about 160 msec by mental preparation and effort.

Latency period could well be an important indicator of cognitive processes, because this period would involve various stages such as : auditory stimuli are transmitted through sensory nerves to the auditory areas of the brain and are perceived ; then via various auditory information processing stages they are replaced by motor codes and are transmitted through motor nerves, thus starting the finger movements. During the retention interval, the subjects manage to use the appropriate fingers to match the pitch information in memory, and they repeat the finger movement pattern as many times as possible. With the repetition of finger movements, the movement becomes more accurate and the rate gradually becomes faster. The elaborate fingering is transmitted to the motor area of the brain, and then the intermodal coordination between auditory pitch information and finger movements would be formed. Thus, the pitch information would be encoded and retained by means of the motor representations.

A series of studies on the latency properties of typing (Shaffer, 1973, 1975 ; Shaffer & Hardwick, 1970) has led to the conclusion that output is produced in two stages. It is assumed that the first stage translates its string simultaneously and deposits the response codes in a buffer memory, while the second stage converts response codes iteratively into finger movements. The two stages overlap in time, so that the contents of the buffer can be regarded as a queue of elements with service (response translation) at one end and bulk renewal at the other. The second stage would require supplementary information about the location of the hand following the previous responses.

Rumelhart and Norman (1982) proposed the simulation model of information processing system involved in typing : The input of the model is a string of characters that constitute the text to be typed. The perceptual processes output their identification of the input to a buffer which maintains the information while the appropriate response (key press) schemata are activated and take control of the actions. The output is a sequence of finger movements, either displayed on a visual computer-controlled display as the movement of the hands and fingers over a typewriter keyboard, or as a series of coordinate locations for the relevant body parts.

In Figure 7–9, an example of the fingering of the tapping with the 6–tone series is illustrated. The five steps of bar height from bottom to top correspond to the five fingers from thumb to little finger. The bar width indicates the duration of touching between the tip of the finger and the aluminum board. An overlap between two successive bars indicates that the two fingers were simultaneously touching the aluminum board. Figure 7–9 shows the following : (a) Latency in Experiment 7–2 (interfering melody) was shorter than in Experiment 7–1 (blank–interval condition) ; that is, tapping started earlier in Experiment 7–2 than in Experiment 7–1. (b) During the retention interval, subjects repeated the tapping much faster and more often in Experiment 7–2 than in Experiment 7–1. It appears that, especially in the interference condition, subjects concentrated on moving their fingers to memorize and encode the pitches. (c) Subjects gradually tried to use more appropriate fingers. Tapping with accurate and elaborate fingering was repeated gradually faster and the amount of tapping increased as time passed. It appeared that during the retention interval subjects came to repeat the movement of tapping automatically. When the comparison melody started, subjects reconstructed their memory for the pitches from their finger movements, matching the pitches of the comparison melody one by one.

The tapping data indicated other results. When the comparison series was the same as the standard series, the subjects continued tapping to match tones and taps one by one, until they heard all the tones of the comparison series. Thus subjects carried out exhaustive scanning (Sternberg, 1966). When the comparison series was the different from the standard series ; in the case of Cs (lower panel of Figure 7–9), exhaustive scanning also tended to be performed ; while in the case of Es (upper panel of Figure 7–9), the moment they listened to a different pitch from the standard series subjects tended to stop tapping. That is, they performed self–terminating search (Sternberg, 1966). In other words, when the subjects employed the tapping strategy (in Session 2), they could detect the different pitch more accurate in Es than in Cs from their motor representations (Figure 7–5).

Experiment 7 was designed to investigate whether the finger–tapping would be effective as a pitch encoding strategy. The conclusions based on these experiments were as follows :

First, for subjects who were highly trained in music, employing the external motor encoding strategy of tapping could be more effective than using internal encoding strategies (pitch rehearsal, representations of melodic contour or staff notation, or verbal rehearsal of note names) to encode and retain the pitches of melodies. The finding is similar to the view that motor–enactment encoding is more effective than visual–imagery encoding on free recall performance of words (Engelkamp, 1986, 1988).

Second, the tapping effect became stronger as the melody length and as ISI duration increased. This was especially true where there were interfering melodies or series of note names during the retention interval. While it would be difficult for subjects to use internal encoding strategies in these challenging situation, motor encoding strategy would be stable or robust. Subjects made the effort to rehearse the tapped pattern much faster and more often in order to retain the standard series. During the retention interval, they gradually tried to use more appropriate, accurate, and elaborate fingering (Figure 7–9). While Cohen (1983) suggested that in subject-performed tasks (SPTs), subjects's conscious rehearsal was not observed and there are no effects of elaboration on recall performance, in Experiment 7 motor elaboration and subjects' conscious efforts to rehearse action were observed in the process of motor encoding.

Third, in order to repeat the tapping many times in certain situations, subjects would start the tapping as early as possible, as reflected in the reduction of the latency period. Internal encoding strategies (such as humming, visualizing, and verbalizing) could not be observed by measuring the number of times and the rate of rehearsal, or the duration of latency between the end of the auditory stimulus and the start of rehearsal. On the other hand, the external encoding strategy of tapping could be observed and analyzed by these measures. Therefore, the second and third conclusions are supported by the data.

Fourth, when not employing the tapping strategy, subjects processed contour and pitch as independent features, and each of them could be an effective cue for melody recognition, as found in Experiment 1 (e.g., Deutsch, 1977 ; Dowling, 1971, 1972, 1978 ; Dowling & Fujitani, 1971 ; Kallman & Massaro, 1979 ; Mikumo, 1992d). Subjects could thus use contour as a cue for melody recognition, especially under interference conditions and as melody length increased (e.g., Edworthy, 1985). Moreover, subjects used contour as a cue with short ISI durations, while they used pitch as a cue with long ISI durations (e.g., Dewitt & Crowder, 1986 ; Dowling & Bartlett, 1981). When the external motor encoding strategy of tapping was employed, the finger movement tended to assimilate the cues of contour and pitch as ISI duration increased ; that is, the finger movement served as an effective means of encoding.

Fifth, each subject would have their own individual dominant internal encoding strategy in ordinary conditions without interference. When subjects were instructed to employ the tapping strategy or when they were under conditions where it was difficult to use their own strategies, they could employ the motor encoding strategy effectively on their own. Since some subjects could not employ the motor encoding strategy independently of verbal encoding, it was necessary for them to employ a multi–coding strategy. This finding is similar to the view that action–event memory is not independent of verbal memory (Ohta, 1993).

The results regarding explicit motor representations found in Experiment 7 provide some evidence supporting important views proposed in the levels–of–processing theory. The greater the degree of challenge in terms of melody length, ISI duration and interfering stimuli, the greater the efforts of the subjects to reduce the latency and to make elaborative and rapid finger movements. This finding supports the view that cognitive (conscious) effort is used for encoding (Battig, 1979 ; Jacoby & Craik, 1979 ; Lockhart, Craik, & Jacoby, 1976). As shown in Figure 7–9, it was also found that the finger movements gradually became more appropriate, accurate and rapid during the retention interval. This finding supports the view of a *continuum* of rehearsal operations running from the minimal processing necessary to repeat a word continuously to various types of elaborative processing involving either further enrichment of one item or associative linkage of several items (Craik, 1979), and the view that the encoding process gradually spreads and becomes richer and more elaborative (Anderson & Reder, 1979 ; Craik & Tulving, 1975 ; Lockhart, Craik, & Jacoby, 1976). Craik and Tulving (1975) explained in terms of verbal stimulus that the minimal core encoding can be elaborated by a context of further structural, phonemic, and semantic encodings. Thus, the more the subjects rehearsed elaborative tapping, the greater the recognition performance became (e.g., Rundus, 1971, 1977, 1980 ; Rundus & Atkinson, 1970 ; Rundus, Loftus, & Atkinson, 1970).

Graphic illustration of the finger movements as in Figure 7–9 revealed that the contour of bar lines was similar to the melodic contour of the standard series ; this provides strong evidence that there is auro–motor coordination (with visual mediator). Furthermore, this intermodal analogy would be based on spatio–motor representations of the relationship between the internal representation of the keyboard's spatial properties and the finger movement pattern. The relationship between this spatio–motor representation and visuo–spatial representation was investigated in Experiments 8 and 9.

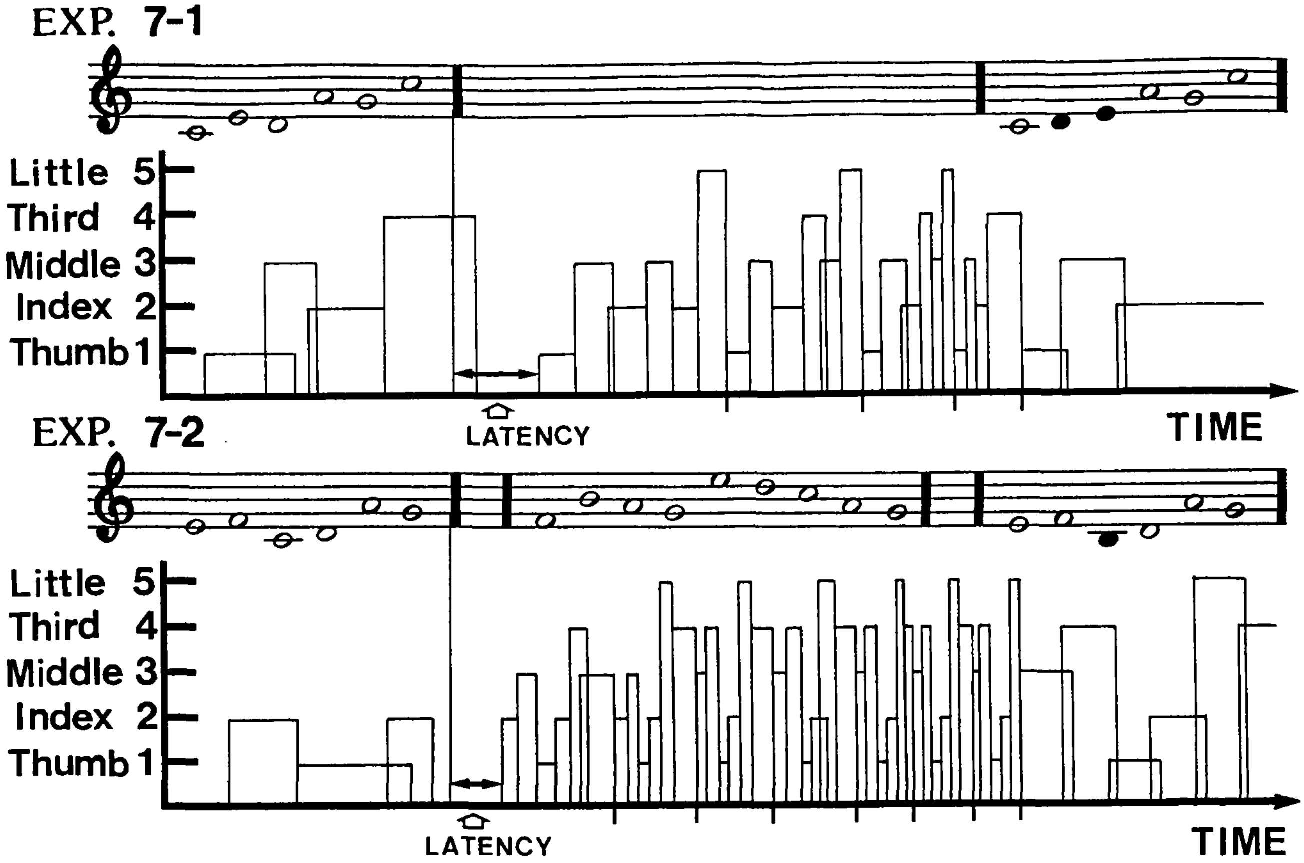


Figure 7–9.

Illustration of the movements of each finger during employment of tapping strategy during ISI in Experiments 7–1 & 7–2. Exhaustive scanning could be seen in Contour–preserving comparison (lower panel), and self–terminating search could be seen in Exchanging comparison (upper panel).

Notes.

Experiment 7–1 : blank interval condition,

Experiment 7–2 : interfering melody condition.

CHAPTER 6

EXPERIMENTAL TESTS OF VISUO-SPATIAL REPRESENTATIONS FOR PITCH

EXPERIMENT 8

Image of Spatial Configurations for Pitch

Method

Results and Discussion

EXPERIMENT 9

Eye Tracking of Visuo-Spatial Representations

Method

Results and Discussion

EXPERIMENT 8

Image of Spatial Configurations for Pitch

When pitch information is encoded, various strategies are employed ; for example, verbal rehearsal of note names, pitch rehearsal through humming or mental rehearsal, visualizing melodic contours or notations, or motor encoding with finger movements analogous to playing the piano, as described in Experiments 1 to 7. Especially in the visualizing or motor encoding strategy, visuo–spatial configurations are used to construct intermodal representations for auditorily presented tones.

Walker (1978, 1981) reported many examples, obtained from children, concerning intermodal analogies between pitch as an auditory stimulus and deployment of visual space. For example, some older children formed an intermodal analogy between pitch and positions along with a vertical ordinate, or between pitch and shape or texture deployments ; a few children matched perceived movements in auditory space with horizontal movements in the visual field ; some younger children matched pitch movements with size differences. Davidson & Scripp (1988) also investigated children's musical intermodal representations.

Pratt (1929) has reported that the composer Berlioz suggests that high and low tones for the pianist lie in the horizontal directions of right and left, and that the violoncellist must reach downward to produce high tones. In Pratt's experiment (1929), the results showed that high tones are phenomenologically higher in space than low ones.

In Experiment 8, relationships between the visuo–spatial configurations for pitch and the motor system used in playing an instrument were investigated (Mikumo, 1993a,d). As de–

scribed in Experiment 7, for the subjects highly trained in playing an instrument, the coordination between auditory pitch information and finger movements has been experienced in relationship to the internal representation of the instrument's spatial properties.

Method

Subjects

Forty–nine subjects who were highly trained in music majoring in piano, violin, violoncello or vocal music, and 17 female subjects who were less well musically trained majoring in literature or domestic science participated in this experiment. Dividing the highly trained subjects by instrument, the number of subjects in the piano, violin, violoncello, and vocal group was 17, 15, 9 and 8, respectively. Their average age, age range, average years of formal training in the specific instrument or vocal music, and the range of years of training, respectively, are shown below :

Group	Age (yeas)	Experience (years)
Group P (Piano)	21.2 (19–23)	16.8 (13–19)
Group Vn (Violin)	22.0 (19–24)	17.1 (13–20)
Group Vc (Cello)	23.6 (19–26)	10.3 (8–15)
Group Vo (Vocal)	22.7 (20–25)	9.7 (7–13)
Group G (General)	20.8 (18–22)	2.2 (0– 4)

In each highly trained group, each subject had had at most five years of formal training in an instrument other than the one indicated. Of the subjects in Group G, none were currently playing a musical instrument.

Materials

Each trial involved a standard series followed by a retention interval with a visual tracking task and then by a comparison series. Of the comparison series, half were exactly the same as the standard series, and the other half were different from the standard. Both the standard and comparison series consisted of eight tones taken from an equal-temperament scale, which ranged from A_3 (220.0 Hz) to E_5 (659.26 Hz). The duration of each tone was 700 msec (600 msec with 100-msec silence), so that the length of the 8-tone series was 5.6 sec. All of the series were tonal melodies in a major key, and were high in tonal melodic structure according to conventional Western rules. They involved a wide variety of contours.

During the retention interval, the visual tracking task was interpolated. The subjects were required to track a circle on a computer display, which appeared successively at 16 points with a slight interval, corresponding to twice the 8-tone standard series. The circle seemed to move with 16 fixed points in each of six directions ; that is, in the X-, Y-, or Z-direction (i.e., left-right, up-down, or backward-forward in perspective), each movement had a positive or negative direction. X-positive direction indicates that the circle moves spatially Rightward (**XR**) corresponding to higher pitch, whereas X-negative direction indicates that the circle moves spatially Leftward (**XL**) corresponding to higher pitch. Y-positive direction indicates that the circle moves spatially Upward (**YU**) corresponding to higher pitch, whereas Y-negative direction indicates that the circle moves spatially Downward (**YD**) corresponding to higher pitch. Z-positive direction indicates that the circle moves Forward in perspective (**ZF**) corresponding to higher pitch, whereas Z-negative direction indicates that the circle moves Backward in perspective (**ZB**) corresponding to higher pitch. There were three types of movements of the circle. "Same" (**S**) indicated that circle movement was exactly the same as that of the 8-tone standard series ; the circle moved with eight fixed points, and the visual distances between the fixed points exactly corresponded to the auditory pitch intervals of the standard series. "Contour-preserving" (C) was obtained by changing the pitch of two tones preserving the melodic contour (the ascending and descending pattern) of the standard series ; the visual distances between the fixed points were somewhat different from the auditory pitch intervals. "Exchanging" (E) was obtained by exchanging the order of two successive tones in the standard series ; the visual distances between the fixed points were completely different from the auditory pitch intervals. Therefore, the circle moving in one of 18 patterns (6 directions X 3 types) on the computer display was the visual stimulus interpolated during the retention interval. The combinations of the 6 directions (XR, XL, YU, YD, ZF, ZB) with the 3 types (S, C, E) are indicated by the terms XR–S, XR–C, XR–E, XL–S, XL–C, XL–E, YU–S, YU–C, YU–E, YD–S, YD–C, YD–E, ZF–S, ZF–C, ZF–E, ZB–S, ZB–C, and ZB–E. In the case of the X– and Y–direction, the circle with a diameter of 1.5 cm moved horizontally or vertically, whereas in the case of the Z–direction, the circle with a diameter of 0.5 cm (backward) to 2.0 cm (forward) moved on a center line between two side lines in perspective. The circle appeared with a constant duration of 500 msec[(5)] on eight fixed points in a series, and the series (4.0 sec) was repeated twice during the retention interval. The interval between the former half and the latter half was also 500 msec. Therefore, each trial involved a 5.6–sec standard series followed by a 2.0–sec interval, an 8.5–sec visual tracking task, a 2.0–sec interval, and then a 5.6–sec comparison series.

The tones of the standard and comparison series were generated by an NEC PC–9801 26K sound synthesizer board installed on an NEC PC–9801 DA personal computer, recorded on tape, and presented over high–quality sound reproduction equipment (DENON digital audio tape recorder DTR–80P).

Procedure

Each subject sat in front of a personal computer and wore headphones, and all tones were adjusted to be equal in loudness. The subjects were instructed that this was an experiment on memory for melodies, and that in each of the 72 trials, they would first see a trial number on a computer display, then hear a first melody (standard series), which would be followed by a visual tracking task in which a circle would move corresponding to twice the 8–tone first melody, and then a second melody (comparison series). The interval between trials was 15 sec. The subjects were required to judge whether the two melodies were the same or different in pitch, and to indicate their judgments by writing "S" (Same) or "D" (Different) on an answer sheet. They were also required to rate the incompatibility experienced when visually tracking the circle movement on a five–point scale, with responses ranging from "no incompatibility (0)" to "very high incompatibility (4)". The subjects were also instructed that they should fixate their eyes on and track the movement of the circle, but that they should not employ any motor encoding strategy analogous to playing the piano. Four practice trials with feedback were given to the subjects prior to the 72 trials.

For this experiment, three lists were prepared and each list involved 72 trials. These lists were counterbalanced between subjects in each group, to minimize the differences among the trials in each of the 18 patterns of visual stimuli in difficulty. In each list, four trials were run for each pattern of visual stimuli. The 18 patterns of visual stimulus were alternated, and the order of presentation of all 72 trials was randomized.

Results and Discussion

Recognition probability data (hit rate plus correct rejection rate) and incompatibility data for 18 visual patterns were calculated in each group.

Recognition probability data for S (Same) in the six directions (XR–S, XL–S, YU–S, YD–S, ZF–S, ZB–S) in each group were analyzed in a one–way analysis of variance, with repeated measure on Pattern factor. When a significant main effect of Pattern factor was obtained, multiple comparisons among the six patterns were analyzed using Newman-Keuls method. In Group P, there was a significant main effect of Pattern [$F(5,80)=5.54$, $p<.001$]. The main effect indicates that the recognition probabilities for YU–S and XR–S were significantly higher than that for XL–S (significance levels between YU–S and XL–S, and XR–S and XL–S were both $p<.01$). In Group Vn, there was a marginal main effect of Pattern [$F(5,70)=2.24$, $p=.059$], but there were no significant differences among these six patterns. In Group Vc, there was a significant main effect of Pattern [$F(5,40)=5.35$, $p<.01$]. The main effect indicates that the recognition probabilities for ZB–S and YD–S were significantly higher than that for XL–S (both $p<.01$). In Groups Vo and G, there was no significant main effect of Pattern.

Incompatibility data for S (Same) in the six directions (XR–S, XL–S, YU–S, YD–S, ZF–S, ZB–S) in each group were analyzed in a one–way analysis of variance with repeated measure on Pattern factor. In Group P, there was a significant main effect of Pattern [$F(5,80)=5.15$, $p<.01$]. The main effect indicates that the incompatibilities for YU–S and XR–S were significantly lower than that for XL–S (significance levels between YU–S and XL–S, and XR–S and XL–S were both $p<.01$). In Group Vn, there was a significant main effect of Pattern [$F(5,70)=5.98$, $p<.001$]. The main effect indicates that the incompatibilities for YU–S and ZF–S were significantly lower than those for XL–S and YD–S (significance levels between YU–S and XL–S, YU–S and YD–S, ZF–S and XL–S, and ZF–S and YD–S were all $p<.01$). In Group Vc, there was a significant main effect of Pattern [$F(5,40)=6.75$, $p<.001$]. The main effect indicates that the incompatibilities for ZB–S and YD–S were significantly lower than that for XL–S (both $p<.01$). In Group Vo, there was a significant main effect of Pattern [$F(5,35)=3.38$, $p<.05$]. The main effect indicates that the incompatibility for XR–S was significantly lower than those for YU–S and ZF–S (both $p<.05$). In Group G, there was a significant main effect of Pattern [$F(5,80)=3.35$, $p<.01$].

The main effect indicates that the incompatibilities for ZB–S and YU–S were significantly lower than that for YD–S (both $p<.05$).

Recognition probability data in the five groups (P, Vn, Vc, Vo, G) for S(Same) in each of six directions were also analyzed in a one-way analysis of variance with repeated measure on Group factor. For Pattern XR–S, there was a significant main effect of Group [$F(4,61)=3.25, p<.05$]. The main effect indicates that the recognition probabilities in Groups Vn and P were significantly higher than that in Group G (significance levels between Groups Vn and G, and Groups P and G were both $p<.05$). For Patterns XL–S and YU–S, there was no significant main effect of Group. For Pattern YD–S, there was a significant main effect of Group [$F(4,61)=4.29, p<.01$]. The main effect indicates that the recognition probabilities in Groups Vo and Vc were significantly higher than that in Group G (both $p<.05$). For Pattern ZF–S, there was a significant main effect of Group [$F(4,61)=7.27$, $p<.001$]. The main effect indicates that the recognition probability in Group Vn was significantly higher than that in Group Vc ($p<.05$) and than that in Group G ($p<.01$). For Pattern ZB–S, there was a significant main effect of Group [$F(4,61)=4.59, p<.01$]. The main effect indicates that the recognition probabilities in Groups Vc and Vo were significantly higher than that in Group G (both $p<.05$).

Incompatibility data in the five groups (P, Vn, Vc, Vo, G) for S (Same) in each direction were analyzed in a one-way analysis of variance with repeated measure on Group factor. For Pattern XR–S, there was a significant main effect of Group [$F(4,61)=4.54, p<.01$]. The main effect indicates that the recognition probabilities in Groups Vo and P were significantly lower than that in Group G (significance levels between Groups Vo and G, and Groups P and G were both $p<.01$). For Pattern XL–S, there was a significant main effect of Group [$F(4,61)=2.58, p<.05$], but there were no significant differences among the five groups. For Pattern YU–S, there was a significant main effect of Group [$F(4,61)=4.44$,

$p<.01$]. The main effect indicates that the recognition probabilities in Groups Vn and P were significantly lower than that in Group G (both $p<.01$). For Pattern YD–S, there was a significant main effect of Group [$F(4,61)=6.39$, $p<.001$]. The main effect indicates that the recognition probabilities in Groups Vo and Vc were significantly lower than that in Group G (both $p<.01$). For Pattern ZF–S, there was a significant main effect of Group [$F(4,61)=3.55$, $p<.05$]. The main effect indicates that the recognition probability in Group Vn was significantly lower than that in Group G ($p<.01$). For Pattern ZB–S, there was a significant main effect of Group [$F(4,61)=4.17$, $p<.01$]. The main effect indicates that the recognition probability in Group Vc was significantly lower than those in Groups G and Vn (both $p<.05$).

When there was high recognition probability and low incompatibility in Pattern YU–S, for example, it would indicate that performance was facilitated by visual tracking of the circle which moved vertically upward corresponding to higher pitch, preserving exactly the same distance between fixed points (i.e., circles) as the pitch intervals of the standard melody ; that is, the subject would have an image of spatial configuration in which the upper direction corresponded to higher pitch. In Pattern XL–S, when there was low recognition probability and high incompatibility, it would indicate that performance was disrupted by visual tracing of the circle which moved horizontally leftward corresponding to higher pitch ; that is, the subject would not have a spatial image showing the left direction as corresponding to higher pitch.

From this, the common results derived from those four types of ANOVAs raise the following suggestions (Figure 8–1, Table 5) : For piano majors, the performance was significantly facilitated by XR–S and YU–S, and it was significantly disrupted by XL–S. Therefore, they would have spatial images in which the right and upper directions are compatible with higher pitch, while the left direction is incompatible with higher pitch.

For violin majors, the performance was significantly facilitated by ZF–S, and it was significantly disrupted by XL–S and YD–S. Therefore, they would have a spatial image in which the forward direction is compatible with higher pitch, while the lower and left directions are incompatible. For violoncello majors, the performance was significantly facilitated by YD–S and ZB–S, and it was significantly disrupted by XL–S. Therefore, they would have spatial images in which the lower and backward directions are compatible with higher pitch, while the left direction is incompatible. For vocal music majors, the performance was significantly facilitated by XR–S, YD–S and ZB–S, and it was not significantly disrupted by any pattern. Therefore, they would have spatial images in which the right, lower and backward directions are compatible with higher pitch. For the subjects in Group G, the performance was significantly facilitated by YU–S, and it was significantly disrupted by YD–S. Therefore, they would have a spatial image in which the upper direction is compatible with higher pitch, while the lower direction is incompatible. In Figure 8–1, the black bars indicate the compatible directions and the dotted–line bars indicate the incompatible directions in each group.

Table 5

Compatible and incompatible spatial directions with higher pitch for each group

Group		**Compatible direction(S)**	**Incompatible direction(s)**
Group P	(Piano)	XR YU	XL
Group Vn	(Violin)	ZF	XL YD
Group Vc	(Cello)	YD ZB	XL
Group Vo	(Vocal)	XR YD ZB	
Group G	(General)	YU	YD

There are some relationships between the motor system in playing an instrument and the spatial configuration for pitch. On the keyboard of a piano, higher pitch keys are located on the right, and lower pitch keys on the left. On the finger–board of a violin, higher pitch fingering positions are located in a forward direction, and lower pitch fingering positions in a backward direction. The violin majors reported that the two side lines of ZF stimuli seemed like the finger–board in playing the violin in perspective. On the finger–board of a violoncello, higher pitch fingering positions are located in a lower direction, and lower pitch fingering positions in an upper direction. The violoncello majors reported that, in playing a violoncello, close–forward fingering positions produce low pitch, and far–backward fingering positions high pitch. Therefore, the movement of the circle in the close–forward direction corresponding to lower pitch and that in the far–backward direction to higher pitch in ZB stimuli in perspective felt highly compatible. Moreover, they felt high compatibility between low pitch and large circles in the forward position, and between high pitch and small circles in the backward position in ZB stimuli. The performance of the vocal music majors showed a tendency similar to that of the violoncello majors. The vocal music majors reported that when they sing in high pitch, they attempt to draw in the chin, to stretch the back of the neck, and to look toward positions far below. For example, when standing on a stage and singing in high pitch, they look toward the back doors of the concert hall ; they reported that the two side lines of ZB stimuli in perspective seemed like the side walls of a concert hall. In general, many people without special training in playing an instrument might innately have a spatial image in which the upper direction corresponds to higher pitch. One reason for this might be that most people have experiences with notes represented on a staff since their childhood in school.

Recognition probability data and incompatibility data for each of the 18 visual patterns in the five groups are shown in Figure 8–2, and the differences in incompatibility between C and S, and between E and S calculated in six directions are shown in Figure 8–3. In Figures 8–2 and 8–3, the black bars indicate the compatible directions and the dotted–line bars indicate the incompatible directions in each group. As described above, C indicates that the circle moved with eight fixed points and two of them somewhat deviated from the exact positions of the standard melody. E indicates that the order of two successive points of eight fixed points was exchanged, and that these two points completely deviated from the exact positions of the standard melody. Therefore, the incompatibility data for C minus S and for E minus S in each direction would become an indicator of the detective power for the deviation between the visual distances of fixed points (i.e., circles) and the auditory pitch intervals. High incompatibility for C minus S or for E minus S would indicate high detective power for the deviation in each direction.

As shown in Figure 8–3, it was found that the detective power for E was consistently higher than that for C. For the piano majors, the detective power for C and E was significantly higher in the compatible spatial directions of XR and YU than in the incompatible direction of XL (significance levels between XR–C and XL–C, XR–E and XL–E, YU–C and XL–C, and YU–E and XL–E were all $p<.01$). For the violin majors, the detective power for C and E was significantly higher in the compatible direction of ZF than in the incompatible directions of XL and YD (significance levels between ZF–C and XL–C, and ZF–C and YD–C were $p<.05$; between ZF–E and XL–E, and ZF–E and YD–E were $p<.01$). For the violoncello majors, the detective power for C and E was significantly higher in the compatible directions of YD and ZB than in the incompatible direction of XL (significance levels between ZB–C and XL–C was $p<.05$; between YD–E and XL–E, and ZB–E and XL–E were $p<.01$) . For the subjects in Group G, the detective power for C and E was somewhat but not significantly higher in the compatible direction of YU than in the incompatible direction of YD.

It was found that the detective power for the deviation between visual distances of fixed points (i.e., circles) and auditory pitch intervals was higher in spatial directions compatible with pitch than in the incompatible directions. When pitches in melodies were encoded, the visuo-spatial representations for pitch information were considerably more accurate in the compatible spatial directions related to the motor system used in playing the instrument. The purpose of Experiment 9 was to investigate whether eye movements reflect the process involved in rehearsing or maintaining the visuo-spatial representations.

****

Footnote (5) —— see Experiment 4.

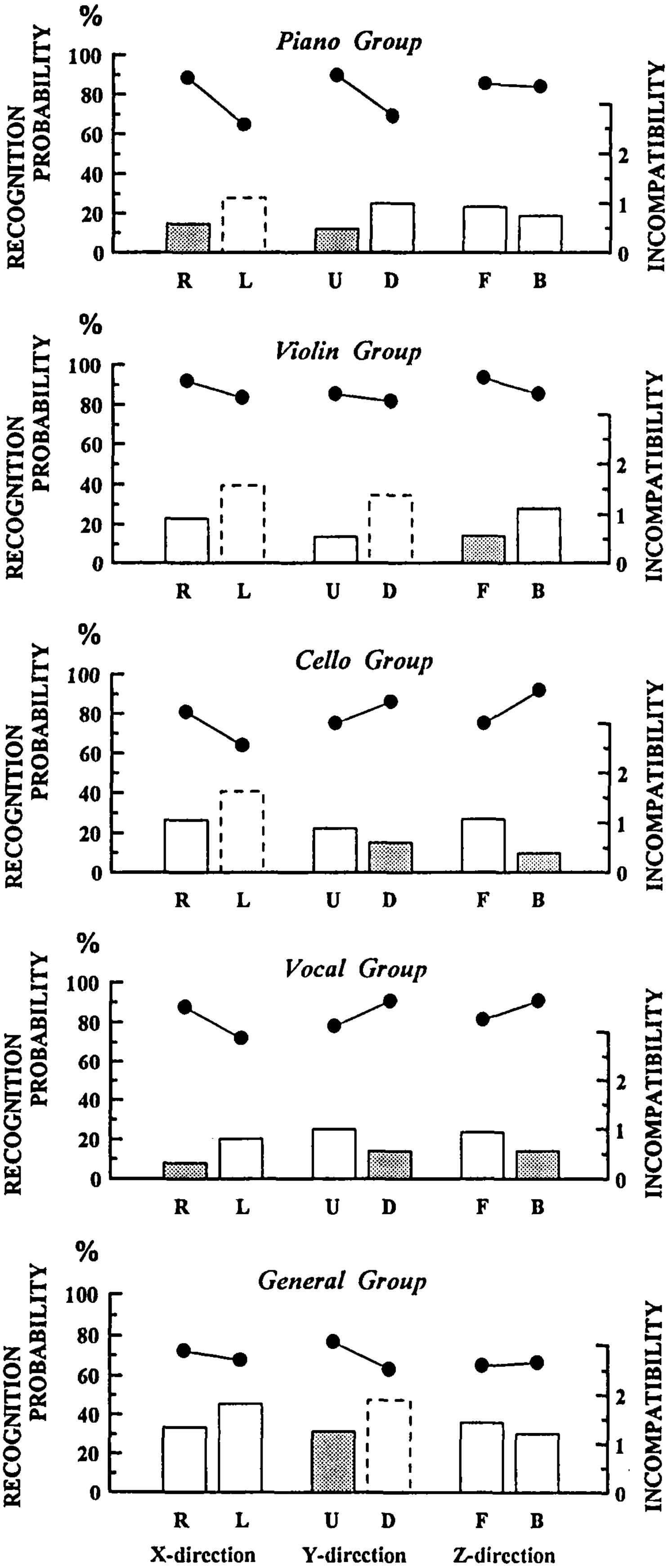


Figure 8–1.

Mean probability of correct recognition (hit rate plus correct rejection rate) and Incompatibility for the S type in the six directions in the five groups. Circles indicate recognition probability and bars indicate incompatibility. Black bars indicate the compatible directions and dotted–line bars indicate the incompatible directions in each group.

Notes.

R : Rightward	U : Upward	F : Forward
L : Leftward	D : Downward	B : Backward

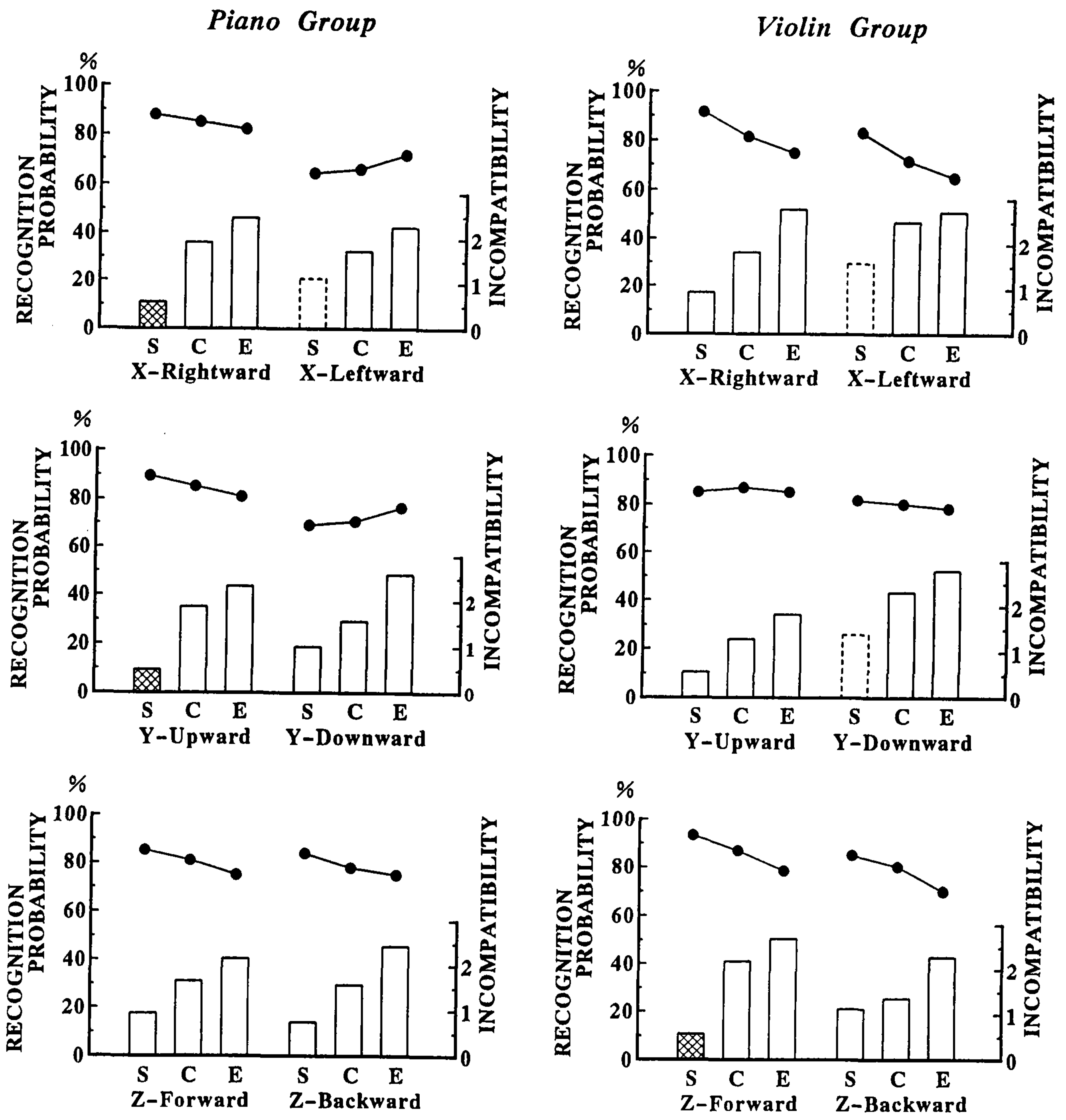
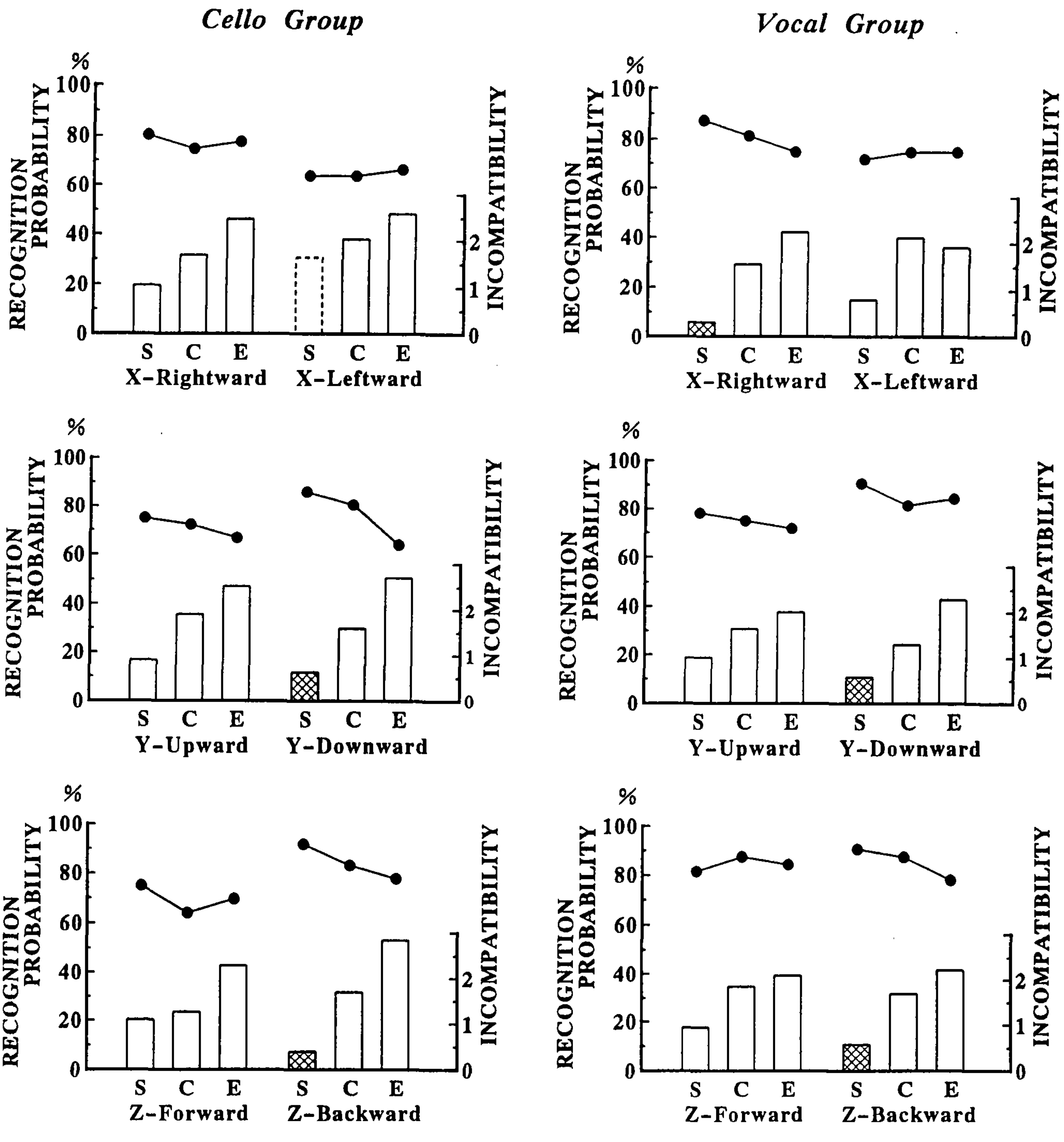


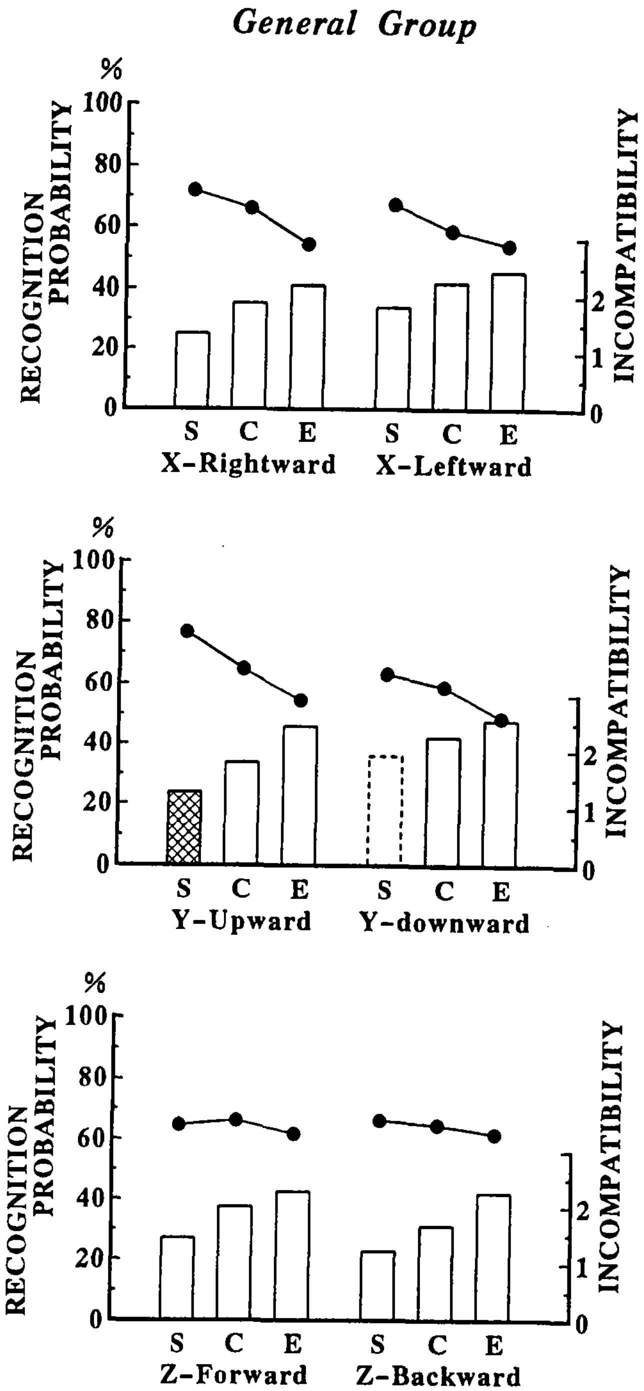
Figure 8-2.

Mean probability of correct recognition (hit rate plus correct rejection rate) and Incompatibility for the three types (S, C, E) in the six directions in the five groups. Circles indicate recognition probability and bars indicate incompatibility. Black bars indicate the compatible directions and dotted-line bars indicate the incompatible directions in each group.

Notes.

S : Same C : Contour-preserving E : Exchanging


Cello Group
Vocal Group
%
RECOGNITION PROBABILITY
INCOMPATIBILITY
100
80
60
40
20
0
2
1
0
S C E S C E
X-Rightward X-Leftward
Y-Upward Y-Downward
Z-Forward Z-Backward


General Group
%
RECOGNITION PROBABILITY
INCOMPATIBILITY
100
80
60
40
20
0
2
1
0
S C E
X-Rightward
S C E
X-Leftward
Y-Upward
Y-downward
Z-Forward
Z-Backward

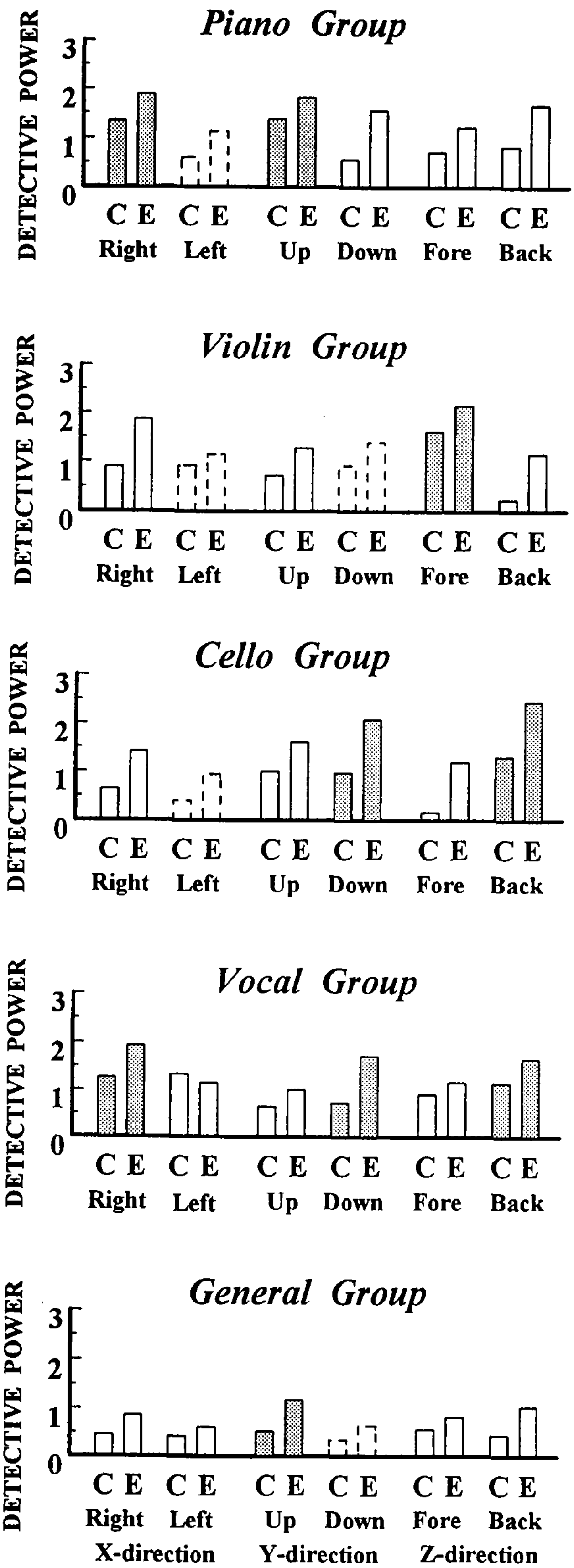


Figure 8–3.

Mean detective power for the two types (C, E) in six directions in the five groups. Black bars indicate the compatible directions and dotted–line bars indicate the incompatible directions in each group.

Notes.

Detective power for C = Incompatibility for C minus S in each direction.

Detective power for E = Incompatibility for E minus S in each direction.

(S : Same, C : Contour–preserving, E : Exchanging)

EXPERIMENT 9

Eye Tracking of Visuo-Spatial Representations

In Experiment 8, the results showed that when pitches are encoded internal visuo-spatial representations are utilized for pitch information, and that these representations are more accurate in compatible spatial directions related to the motor system used in playing the instrument. As in Experiment 5, the purpose in Experiment 9 was to investigate whether eye movements reflect the process involved in rehearsing or maintaining visuo-spatial representations to encode pitch information. This experiment was focused on eye movements specifically in the compatible and incompatible spatial directions for pitch in each group found in Experiment 8 (Mikumo, 1994a,d).

In Experiment 9, in which the method was almost the same as that in Experiment 8, the eye movements during the retention interval were recorded using an eyemark recorder. The visual tracking task interpolated during the retention interval was prepared as the external representation analogous to tracking the visuo-spatial image which would be constructed by the auditory information of the standard memory. As in Experiment 5, subjects were required to make a recognition judgment of melodies and to track the visual task. The visual tracking task included some lure patterns, therefore, it was hypothesized that the subject's involuntary matching between the visual task-tracking and mental image-tracking would occur during the retention interval. Thus, accurate visual tracking in response to the correct visual patterns and deviated or confused visual tracking in response to the lure visual patterns in a direction(3) would imply that the direction was compatible with pitch, and that the subjects had retained accurate visuo-spatial representations for pitch information and precisely tracked their visuo-spatial image while encoding pitches. On the other hand, deviated or confused visual tracking in response to even the correct visual patterns in a direction would imply that the direction was inconsistent with subject's visuo-spatial representations for pitch.

Method

Subjects

Twenty subjects who were highly trained in music majoring in piano, violin, violoncello or vocal music (five subjects for each instrument), and five female subjects who were less well musically trained majoring in domestic science participated in this experiment. Their average age, age range, average years of formal training in the specific instrument or vocal music, and the range of years of training, respectively, are shown below :

Group		Age (yeas)	Experience (years)
Group P	(Piano)	20.8 (19–23)	15.3 (13–19)
Group Vn	(Violin)	22.2 (19–25)	16.9 (15–20)
Group Vc	(Cello)	23.4 (21–26)	11.3 (10–14)
Group Vo	(Vocal)	22.3 (20–24)	9.2 (7–12)
Group G	(General)	21.2 (21–23)	2.8 (0– 5)

In each highly trained group, each subject had had at most five years of formal training in an instrument other than the one indicated. Of the subjects in Group G, none were currently playing a musical instrument. All subjects had good eyesight without glasses or contact lenses. None of the subjects were those employed in Experiment 8.

Materials

Each trial involved a standard series followed by a retention interval with a visual tracking task and then by a comparison series. Of the comparison series, half were exactly the same as the standard series, and the other half were different from the standard. During the retention interval, the subjects were required to track a circle on a computer display, which appeared successively at 16 points with a slight interval, corresponding to twice the 8-tone standard series. The circle seemed to move with 16 fixed points in a pattern, and it moved in one of 18 patterns [6 directions (XR, XL, YU, YD, ZF, ZB) X 3 types (S, C, E)] (see Experiment 8). This experiment involved 18 trials corresponding to each of the 18 patterns. The order of presentation of all 18 trials was randomized. In the case of the X-and Y-direction, the circle with a diameter of 1.5 cm moved horizontally or vertically, whereas in the case of the Z-direction, the circle with a diameter of 0.5 cm (backward) to 2.0 cm (forward) moved on a center line between two side lines in perspective. The circle appeared with a constant duration of 500 msec(5) at eight points in a series, and the series (4.0 sec) was repeated twice during the retention interval. The interval between the former half and the latter half was also 500 msec.

Procedure

Each subject sat in front of a personal computer and wore the head unit(6) of an eye mark recorder (NAC Eye Mark Recorder model EMR-600(7)) ; the distance between the computer display and the elliptical mirror unit was about 80 cm. All tones were adjusted to be equal in loudness (approximately 50 dB SPL), and were presented via a loudspeaker. The subjects were instructed that this was an experiment on memory for melodies, and that in each of the 18 trials, they would first see a trial number on a computer display, then hear a first melody (standard series), which would be followed by a visual tracking task in which a circle would appear at 16 points corresponding to twice the 8-tone first melody, and then a second melody (comparison series). The interval between trials was 15 sec. The subject's task was to judge whether the two melodies were the same or different in pitch and to answer orally. The subjects were also instructed that they should track the circle movements, and that they should not employ any motor encoding strategy analogous to playing the piano. After fine adjustments were made to the head unit, the eyeball movement characteristics for the individual subject were calibrated(8). Three practice trials with feed-back were given to the subjects prior to the 18 trials. The other aspects of the method were the same as those in Experiment 8.

Results and Discussion

In this experiment, the subjects were to judge whether the standard series and the comparison series were the same or different in pitch ; however, the purpose of this experiment was not to analyzed the recognition probability, but to investigate the eye movements during the retention interval.

In the 18 patterns [6 directions (XR, XL, YU, YD, ZF, ZB) X 3 types (S, C, E)], the eye-fixation duration on each of the 16 points where the circle appeared in the pattern corresponding to twice the 8-tone standard series were measured. In each pattern, the fixation durations on eight points in the latter half were available, because at the early stage of the retention interval, the subjects' eye movements were not stable. In Figure 9, the average duration of fixation as a function of the serial position in the latter half of each pattern is shown in each group. The data were obtained from the subject's dominant eye.

The visual tracking task of circle movements was interpolated during the retention interval, while the subjects attempted to encode pitches of the standard series. When the circle appeared at eight compatible points which were consistent with the subject's internal visuo-spatial representations, it was hypothesized that it would be easier to track and fixate them, therefore, the fixation durations on these points would be longer, and there would be not marked differences among them. The circle appeared with a constant duration of 500 msec at each point, so that the fixation duration on each point would be 500 msec at most. On the other hand, when the circle appeared at some incompatible points which were inconsistent with the subject's internal visuo-spatial representations, it was hypothesized that it would be difficult to track and fixate them for a long duration, therefore, the fixation durations would be rather shorter on the unexpected points than on the others.

For piano majors, the compatible spatial directions for pitch found in Experiment 8 were the XR and YU directions. In control trials, in which a circle appeared successively at eight points with a constant duration of 500 msec without standard and comparison series, the fixation duration on each point was about 240 msec when the circle moved horizontally (X-direction) and about 270 msec when it moved vertically (Y-direction).

In XR-S and YU-S, in which the visual distances of points were exactly the same as the auditory pitch intervals, there were not marked differences among the fixation durations on the eight points. Each duration was about 360msec in XR-S and about 390 msec in YU-S, and they were much longer than those in the control trials.

Concerning C and E in these directions, in XR-C, a circle appeared at eight points, and the fourth and sixth points somewhat horizontally deviated from the exact positions of the standard melody. The fixation duration was much shorter on these points than on the others, indicating that it was difficult to fixate the circle which appeared at deviated positions. In XR-E, the sixth and seventh points horizontally exchanged order, so that these two points were completely different from the exact positions of the standard melody. The fixation durations were much shorter on these points than on the others, indicating that the subjects could not track well the circle which appeared at these two deviated positions. In YU–C, the fourth and sixth points somewhat vertically deviated, and the fixation durations were shorter on these points than on the others. In YU–E, the third and fourth points vertically exchanged order, and the fixation durations were much shorter on these points than on the others. As in XR–C, XR–E, YU–C and YU–E, the fixation durations were markedly shorter on the unexpected points than on the others, indicating that there was high detective power for the deviation between visual distances of points and auditory pitch intervals in the XR and YU directions.

On the other hand, in the incompatible spatial direction of XL found in Experiment 8, it was difficult to fixate the circle for a long duration even in XL–S, in which the visual distances of points were exactly the same as the auditory pitch intervals. The fixation duration on each point was about 240 msec, and it was almost the same as that in the control trial. In XL–C and XL–E, the fixation durations were short on all eight points, and each was also about 240 msec, indicating that the subjects could not track well not only the circle which appeared at deviated positions but also the circle which appeared at accurate positions. This would indicate that there was not detective power for the deviation between visual distances of points and auditory pitch intervals in the XL direction.

For violin majors, the compatible spatial direction for pitch found in Experiment 8 was the ZF direction. In a control trial, in which a circle appeared successively at eight points with a constant duration of 500 msec without standard and comparison series, the fixation duration on each point was about 240 msec when the circle moved forward–backward (Z–direction).

In ZF–S, there were not marked differences among the fixation durations on the eight points. Each duration was about 350 msec and was longer than that in the control trial. In

ZF–C, in which the fifth and seventh points somewhat deviated from the exact positions, and in ZF–E, in which the fourth and fifth points exchanged order, the fixation durations were much shorter on these points than on the others. As in ZF–C and ZF–E, the subjects could not track well, especially the circle which appeared at unexpected points, indicating that there was high detective power for the deviation between visual distances of points and auditory pitch intervals in the ZF direction.

On the other hand, the incompatible spatial directions found in Experiment 8 were the XL and YD directions. In control trials, the fixation duration on each point was about 220 msec when the circle moved horizontally (X–direction) and about 270 msec when it moved vertically (Y–direction).

Even in XL–S or YD–S, it was difficult to fixate the circle for a long duration. The fixation duration on each point was about 250 msec in XL–S and YD–S and was almost the same as those in the control trials. In XL–C, XL–E, YD–C and YD–E, the fixation durations were short on all eight points, and each duration was also about 250 msec, indicating that the subjects could not track well not only the circle which appeared at deviated positions but also the circle which appeared at accurate positions. This would indicate that there was not detective power for the deviation between visual distances of points and auditory pitch intervals in the XL and YD directions.

For violoncello majors, the compatible spatial directions for pitch found in Experiment 8 were the ZB and YD directions. In control trials, in which a circle appeared successively at eight points with a constant duration of 500 msec without standard and comparison series, the fixation duration on each point was about 240 msec when the circle moved forward–backward (Z–direction) and about 260 msec when it moved vertically (Y–direction).

In ZB–S and YD–S, there were not marked differences among the fixation durations on the eight points. Each duration was about 340 msec in ZB–S and YD–S and was longer than those in the control trials. In ZB–C, in which the third and sixth points somewhat deviated from the exact positions, and in ZB–E, in which the fourth and fifth points exchanged order, the fixation durations were much shorter on these points than on the others. In YD–C, in which the fourth and seventh points somewhat vertically deviated, and in YD–E in which the fourth and fifth points vertically exchanged order, the fixation durations were much shorter on these points than on the others. As in ZB–C, ZB–E, YD–C and YD–E, the subjects could not track well, especially the circle which appeared at unexpected points, indicating that there was high detective power for the deviation between visual distances of points and auditory pitch intervals in the ZB and YD directions.

On the other hand, the incompatible spatial direction found in Experiment 8 was the XL direction. In a control trial, the fixation duration on each point was about 240 msec when the circle moved horizontally (X–direction).

Even in XL–S, it was difficult to fixate the circle for a long duration. The fixation duration on each point was about 250 msec and was almost the same as that in the control trials. In XL–C and XL–E, the fixation durations were short on all eight points, and each duration was also about 250 msec, indicating that the subjects could not track well not only the circle which appeared at deviated positions but also the circle which appeared at accurate positions. This would indicate that there was not detective power for the deviation between visual distances of points and auditory pitch intervals in the XL direction.

For vocal music majors, the compatible spatial directions for pitch found in Experiment 8 were the XR, YD and ZB directions. In control trials, in which a circle appeared successively at eight points with a constant duration of 500 msec without standard and comparison series, the fixation duration on each point was about 220 msec when the circle moved horizontally (X–direction), about 260 msec when it moved vertically (Y–direction), and about 230 msec when it moved forward–backward (Z–direction).

In XR–S, YD–S and ZB–S, there were not marked differences among the fixation durations on the eight points. Each duration was about 330 msec in XR–S and YD–S and about 340 msec in ZB–S, and these durations were longer than those in the control trials. In XR–C, in which the fourth and sixth points somewhat horizontally deviated from the exact positions, and in XR–E, in which the sixth and seventh points horizontally exchanged order, the fixation durations were much shorter on these points than on the others. In YD–C, in which the fourth and seventh points somewhat vertically deviated, and in YD–E, in which the fourth and fifth points vertically exchanged order, the fixation durations were shorter on these points than on the others. In ZB–C, in which the third and sixth points somewhat deviated, and in ZB–E, in which the fourth and fifth points exchanged order, the fixation durations on these points were shorter than on the others. As in XR–C, XR–E, YD–C, YD–E, ZB–C and ZB–E, the subjects could not track well, especially the circle which appeared at unexpected points, indicating that there was high detective power for the deviation between visual distances of points and auditory pitch intervals in the XR, YD and ZB directions.

For the subjects in Group G, the compatible spatial direction for pitch found in Experiment 8 was the YU direction. In a control trial, in which a circle appeared successively at eight points with a constant duration of 500 msec without standard and comparison series, the fixation duration on each point was about 260 msec when the circle moved vertically (Y–direction).

In YU–S, there were not marked differences among the fixation durations on the eight points. Each duration was about 330 msec and was longer than that in the control trial. In YU–C, in which the fourth and sixth points somewhat vertically deviated from the exact positions, and in YU–E, in which the third and fourth points vertically exchanged order, there were small differences between the fixation durations on these points and on the others. This would indicate that, even in the compatible spatial direction of YU, there was not high detective power for the deviation between visual distances of points and auditory pitch intervals.

On the other hand, in the incompatible spatial direction of YD found in Experiment 8, it was difficult to fixate the circle for a long duration in YD–S, YD–C and YD–E. The fixation duration on each point was about 240 msec and was shorter than that in the control trial. In the YD direction as well, there was not detective power for the deviation between visual distances of points and auditory pitch intervals.

It may be concluded that, especially in highly musically trained subjects, in the compatible spatial directions for pitch found in Experiment 8, in S type there were not marked differences among the fixation durations on the eight points, and the duration was much longer than in the control trial. This would indicate that, in the compatible spatial direction, subjects could visually track well the circle movements which appeared at exact points as the standard series. In C and E types in the compatible directions, the fixation durations were markedly shorter on the inexact points than on the others, indicating that there was high detective power for the deviation between visual distances of points and auditory pitch intervals in the directions. These findings suggest that, while encoding pitches, the subjects had retained accurate visuo–spatial representations for pitch information, and that they precisely visually tracked their visuo–spatial images, which are more accurate in compatible spatial directions related to the motor system used in playing the instrument. The fixation duration data support the findings obtained in Experiment 8.

Furthermore, although none of the subjects employed in this experiment were those employed in Experiment 8, the introspections reported by them were similar to those in the subjects in Experiment 8.

Unlike in Experiment 9, in Experiment 5, during the retention interval, eight circles of a melodic contour or eight notes on a staff notation appeared successively from the left to the right ; the number of circles or notes increased one by one with a constant duration up to eight in a series, then all eight circles or notes once disappeared at once with a slight interval, after which the series was repeated again. Therefore, when the circles or notes appeared at deviated positions from the subject's internal representations, in some cases, especially for the highly musically trained subjects, the subjects showed a response to the novelty of the unexpected points. Thus, the fixation durations on these points were longer on the others, or backward tracking to the unexpected points occurred.

On the other hand, in Experiment 9, a circle appeared at eight points successively with a constant duration ; a circle appeared at one point, then, after a slight interval, the circle appeared another point, so that the circle seemed to move with eight fixed points in a series ; and after 500 msec, the series was repeated again. Therefore, when the circle appeared at deviated positions from the subjects's internal representations, the subjects' eye movements were confused or lured by circle movement which appeared at unexpected positions. Thus, the fixation durations on these points were rather shorter than on the others which appeared at appropriate positions where were consistent with the subject's internal representations.

****

Footnotes (3) and (5) —— see Experiment 4.
Footnotes (6) (7) and (8) —— see Experiment 5.

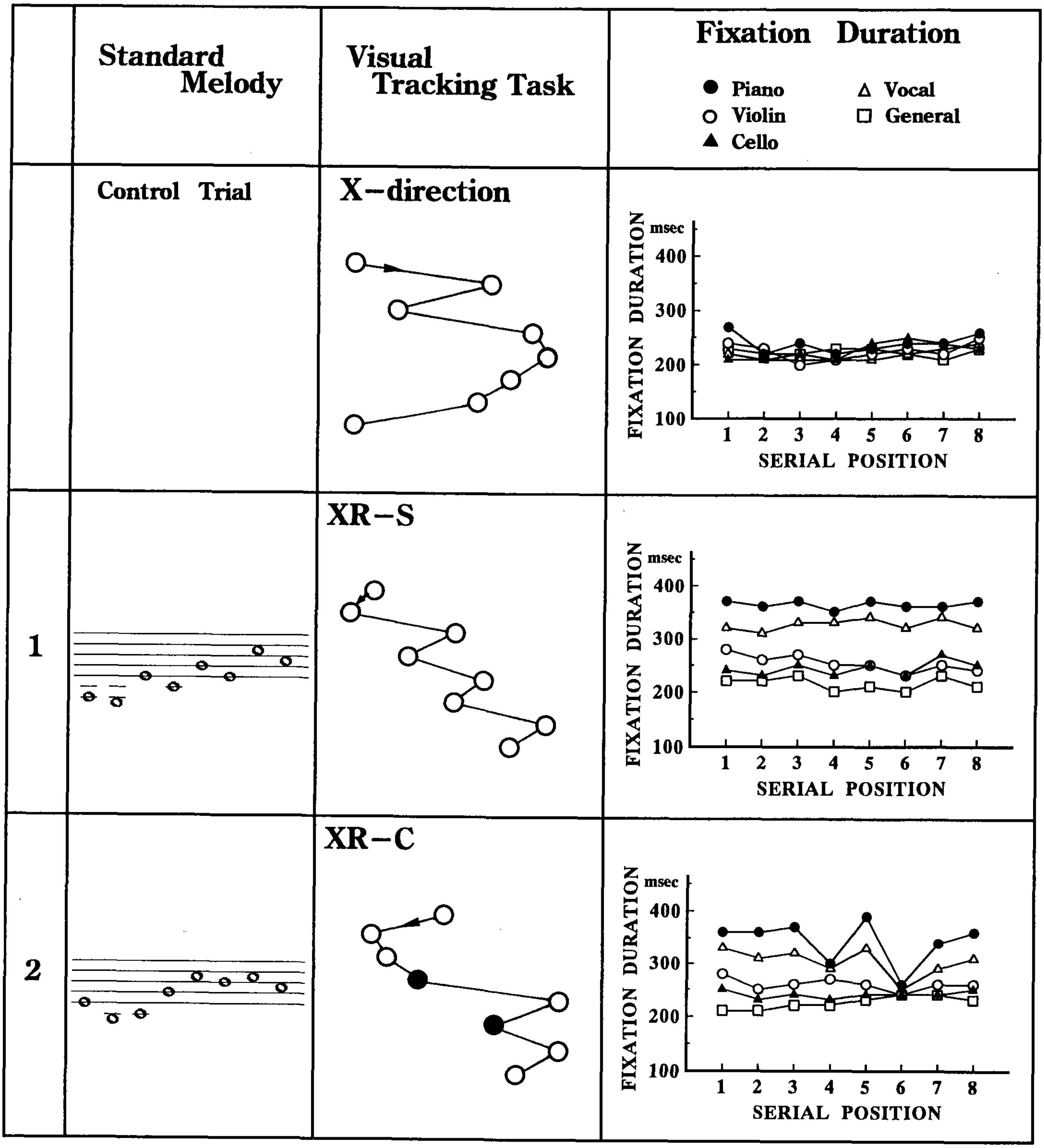
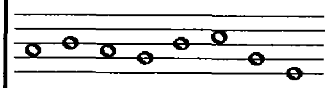
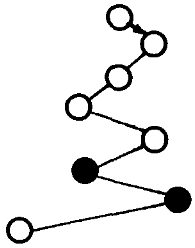
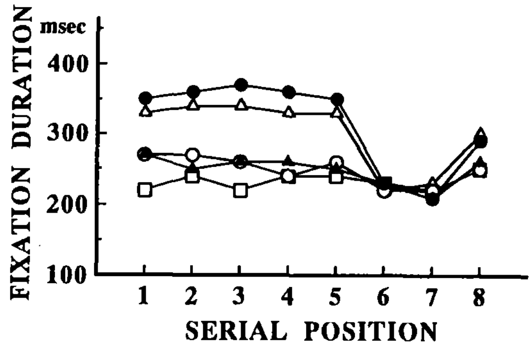
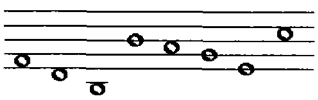
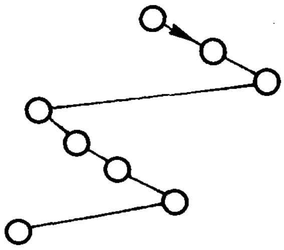
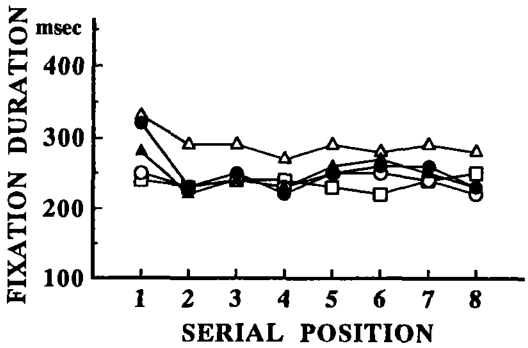
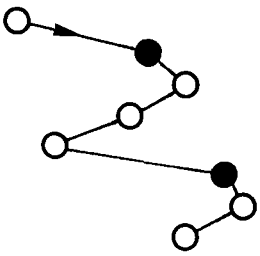
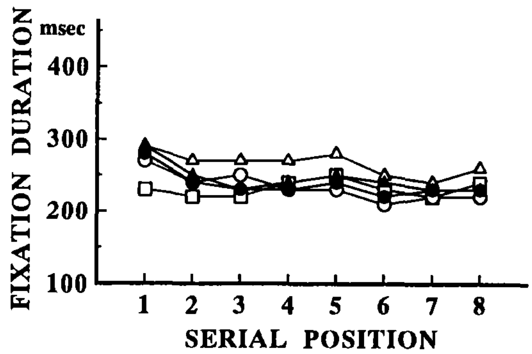
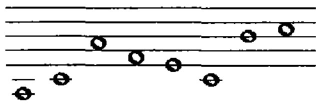
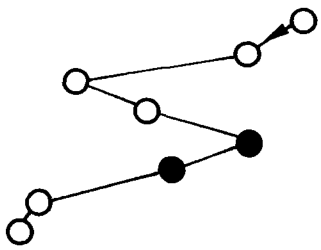
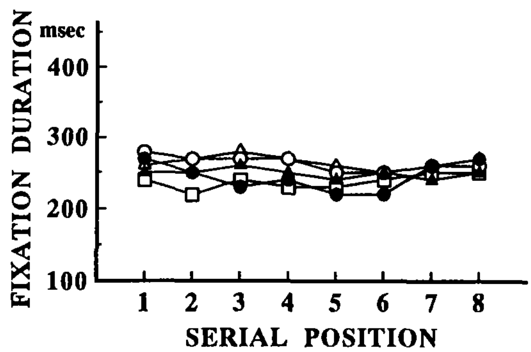
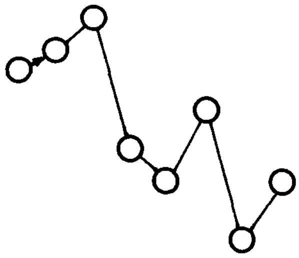
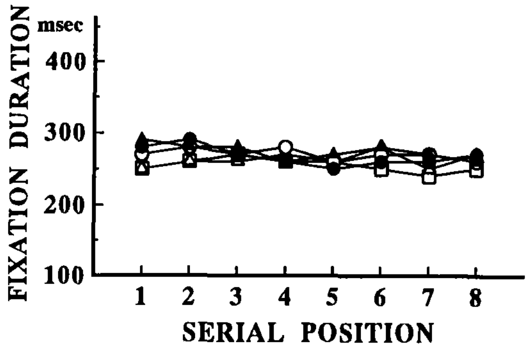
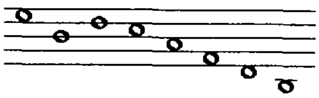
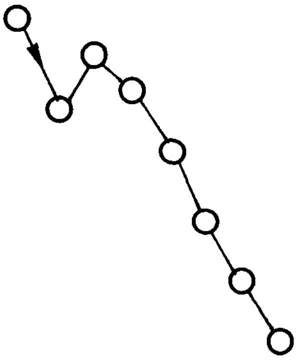
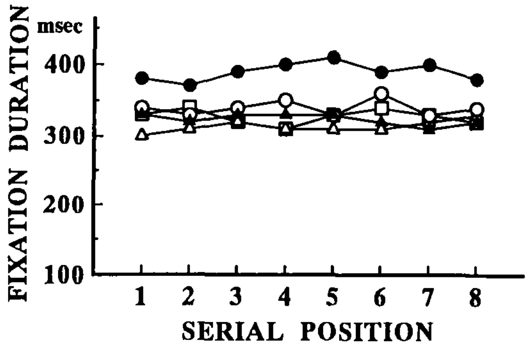
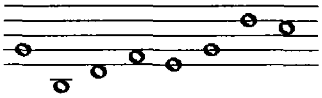
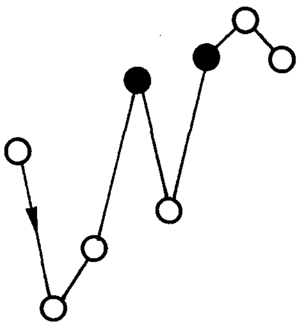
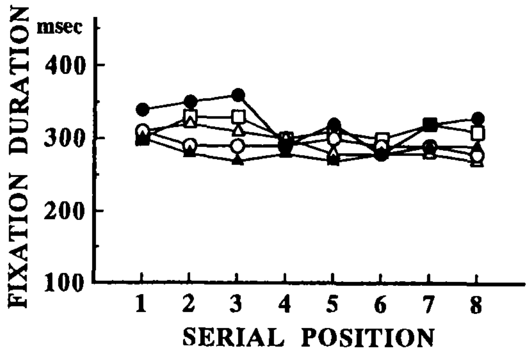
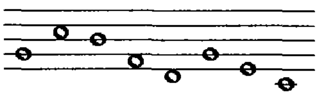
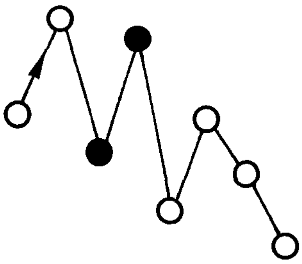
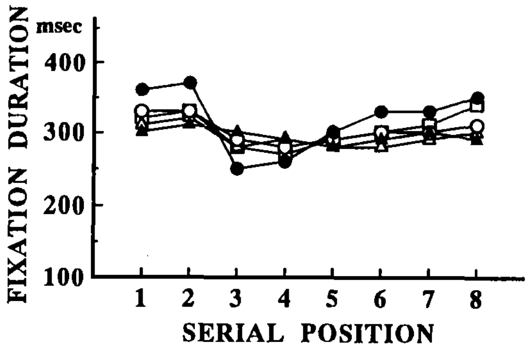
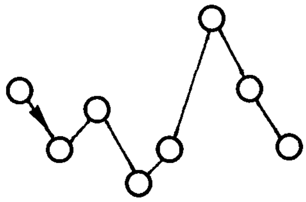
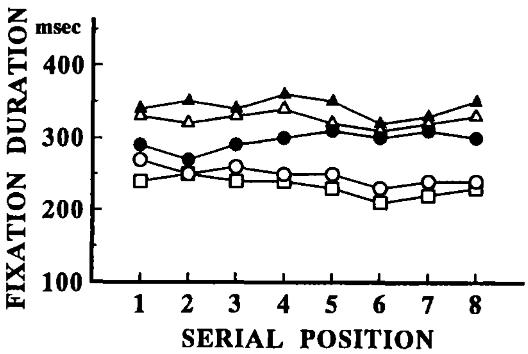
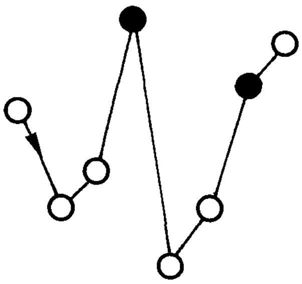
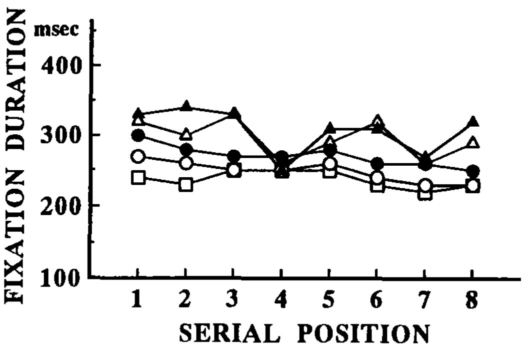
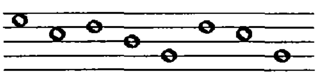
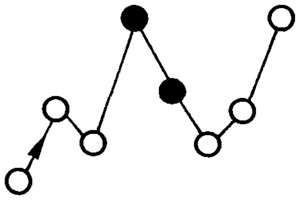
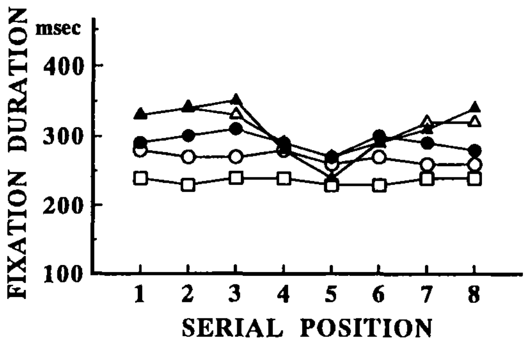


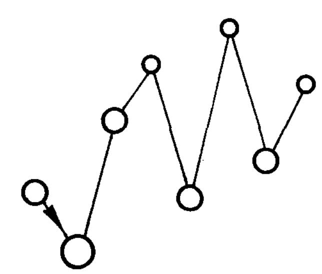
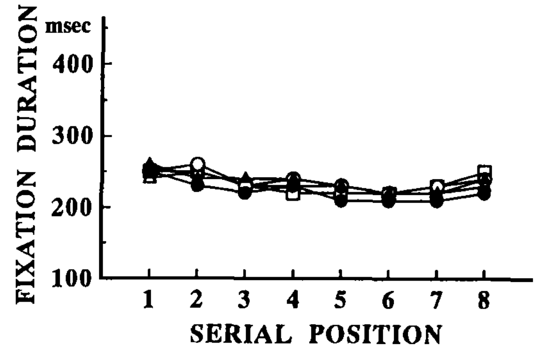
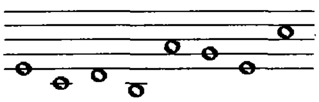
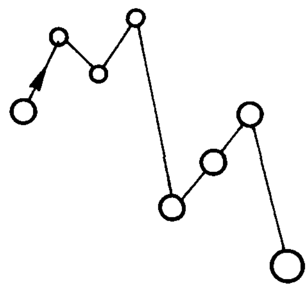
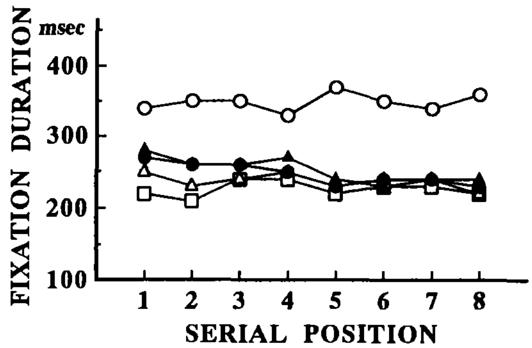
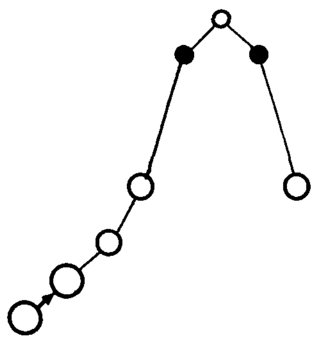
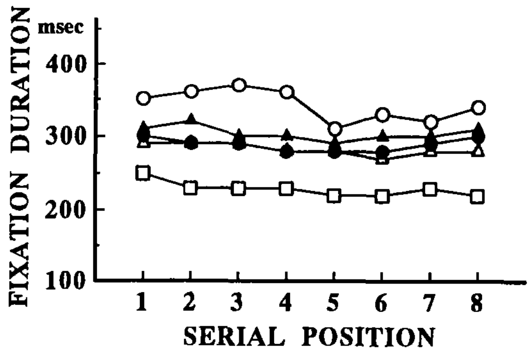
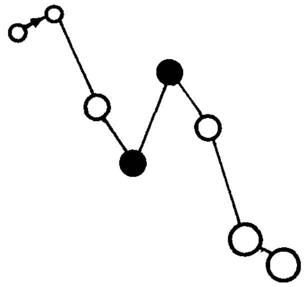
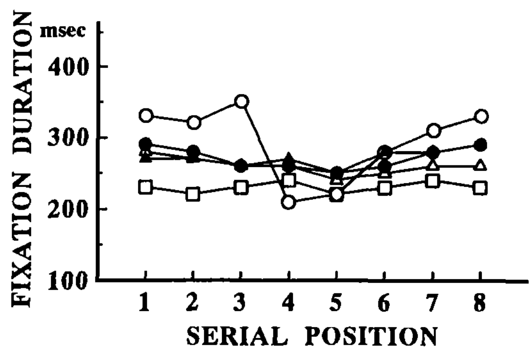
Figure 9.

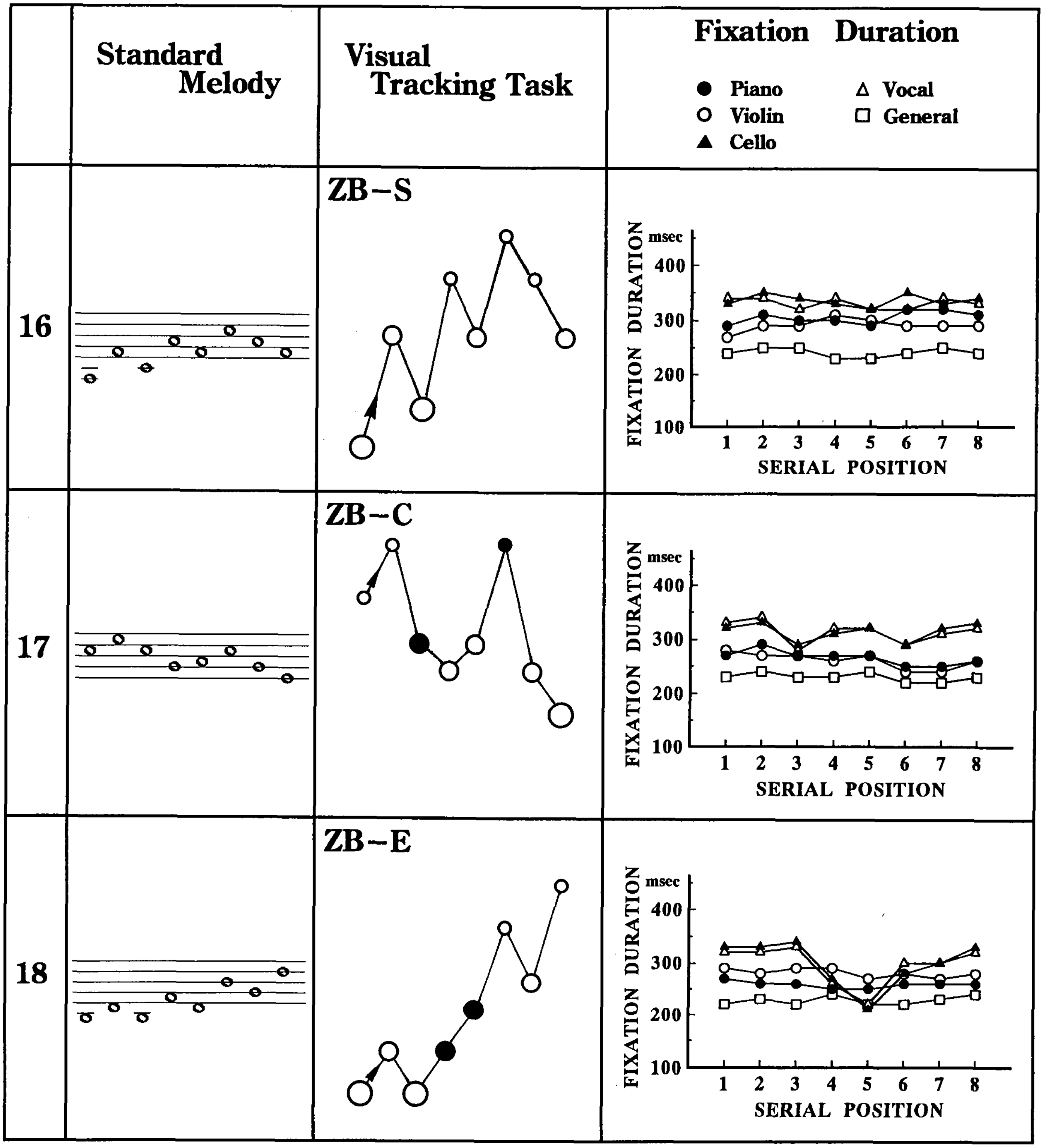
Average duration of fixation for five groups as a function of the serial position in the latter half of each pattern. In the visual tracking task, the circle actually moved on an axis in the X−, Y−axis or Z−axis in perspective. Black circles indicate that their positions are deviated from the exact points.

	Standard Melody	Visual Tracking Task	Fixation Duration ● Piano △ Vocal ○ Violin □ General ▲ Cello
3		XR−E 	
4		XL−S 	
5		XL−C 	
6		XL−E 	

<table>
<tr><th></th><th>Standard Melody</th><th>Visual Tracking Task</th><th>Fixation Duration
● Piano △ Vocal
○ Violin □ General
▲ Cello</th></tr>
<tr><td></td><td>Control Trial</td><td>Y–direction
</td><td>
</td></tr>
<tr><td>7</td><td></td><td>YU–S
</td><td>
</td></tr>
<tr><td>8</td><td></td><td>YU–C
</td><td>
</td></tr>
<tr><td>9</td><td></td><td>YU–E
</td><td>
</td></tr>
</table>

	Standard Melody	Visual Tracking Task	Fixation Duration ● Piano △ Vocal ○ Violin □ General ▲ Cello
10		YD−S 	
11		YD−C 	
12		YD−E 	

	Standard Melody	Visual Tracking Task	Fixation Duration ● Piano △ Vocal ○ Violin □ General ▲ Cello
	Control Trial	Z−direction 	
13		ZF−S 	
14		ZF−C 	
15		ZF−E 	


Standard Melody
Visual Tracking Task
Fixation Duration
● Piano
○ Violin
▲ Cello
△ Vocal
□ General
16
ZB–S
17
ZB–C
18
ZB–E
FIXATION DURATION
msec
400
300
200
100
1 2 3 4 5 6 7 8
SERIAL POSITION

CHAPTER 7

SUMMARY AND CONCLUSIONS

SUMMARY AND CONCLUSIONS

The purpose of this study was to investigate the encoding strategies for pitch information of short melodies in terms of imagery-based multimodal representations. The findings of these experiments are summarized below.

When the subjects were asked to memorize and retain the pitch information in short melodies, the performance of the highly musically trained subjects was consistently better than that of the untrained subjects, because the former subjects could listen analytically to the musical series, with considerable attention to the internal relationships among their components (i.e., pitch interval), and because they employed their own strategy effectively or were motivated to try to employ several strategies at the same time to encode pitch information. Under ordinary conditions, their dominant and effective strategy of encoding the pitches of tonal melodies was verbal rehearsal of note names. When this verbal encoding strategy was employed, pitch or pitch intervals could be retained accurately for a long time (Experiment 1).

The neuropsycological evidence that the highly musically trained subjects use verbal encoding strategy for tonal melodies was found in Experiment 3–3, in which the tonal melodies, which are actually non-verbal stimuli in themselves but are "note name-evoking" stimuli, would be processed as verbal stimuli in the left hemisphere. The preferential left hemisphere lateralization found in the highly musically trained subjects would be due to the cognitive "linguistic" structure of tonal melodies, which implies the likelihood that pitches are encoded as verbal labels (note names) and that the processing of musical tasks involves sequential programs most analogous to those of language and speech. On the other hand, the atonal melodies would be processed as non-verbal stimuli in the right hemisphere. In Experiment 3-4, the stimuli were the melodies sung with note names at accurate pitches, so that the subjects were given the verbal codes at the same time and it was not necessary for them to encode pitches as verbal code (note names) by themselves. In this experiment, the highly musically trained subjects could retain the pitches not only in tonal melodies but also in atonal melodies as verbal codes ; when there was a possibility that the pitches were encoded as note names, these melodies were processed in the left hemisphere. However, the less well musically trained subjects could not encode pitches as note names, even when the pitches in a melody and their verbal codes were given at the same time, and they therefore processed tonal and atonal melodies as non-verbal stimuli in the right hemisphere.

The ability to encode pitch information as verbal labels would be closely related to the ability to detect a 50-cent deviated tone out of a melody, because the 50-cent deviated tone from an equal-temperament scale is the most difficult pitch to categorize into a note name on a chromatic scale. Although the ability depends considerably on experience or training in music, it was found that even those who were less well trained in music had acquired the ability to some extent as their age increased. Moreover, the later the deviated tone appeared in the melody, the greater the accuracy with which it was detected. In the detection of a deviated tone out of a melody, the subject gradually constructs an internal cognitive framework (scale schema) upon hearing the tones from the beginning of the melody. This cognitive framework plays an essential role in the perceptual interpretation of each succeeding tone, and the subjects can detect a deviated pitch out of a melody by referring to the framework. In this experiment, the subjects would constructed the diatonic scale schema upon hearing the tonal melodies, so that the detection of a deviated tone out of the tonal melodies was more accurate than out of the atonal melodies (Experiment 2).

Since it was difficult for the less well musically trained subjects to encode pitches as note names, they attempted to encode pitches as acoustic pitch codes through humming or mental rehearsal of pitches, and this was also true for the highly trained subjects with atonal melodies. In this case, they tended to listen globally to the pitch sequences on the basis of the total configuration ; the melodic contour was the dominant and effective cue for encoding pitch sequence (Experiment 1). Melodic contour is the other property as important as pitch or pitch interval in terms of melody recognition, and there is an assumption that various visual and auditory contours are perceived in broadly similar ways.

In Experiment 4, the auro–visual intermodal relationships based on visual representations were investigated. It was found that the less well musically trained subjects employed a visualizing strategy in which an auditory contour (auditory imagery) was visualized as a visual contour (visual imagery), and that visualization of auditory imagery as melodic contours would be to some extent an effective strategy for them to retain pitch sequences. Their internal visual representations of melodic contour roughly reflect the auditory pitch intervals. In stead of melodic contour, the highly musically trained subjects visualized pitch information as staff notations, in which notes appeared at accurate positions on a staff, especially for tonal melodies. That is, the visual distance between notes in their internal representations precisely reflect the auditory pitch intervals. Visualization of auditory imagery as staff notations would be an effective strategy to encode pitches of melodies.

The evidence that the highly musically trained subjects precisely tracked their internal visual representations of notes on a staff, especially for tonal melodies, and that the less well trained subjects roughly tracked their internal visual representations of melodic contour while employing visualizing strategy, was obtained in the study of eye movements (Experiment 5).

In Experiment 6, it was found that subjects employ two or three codes at the same time rather than just one to memorize or retain pitches. For example, pitch rehearsal of auditory information along with note names (dual coding) and, to a greater extent, at the same time visualizing the staff notation (triple coding) were the most effective strategies for the highly musically trained subjects with tonal melodies. Pitch rehearsal along with visualizing the melodic contour (dual coding) was effective strategy for the highly musically trained subjects with atonal melodies and for the less well trained subjects with both tonal and atonal melodies.

In Experiment 7, the intermodal auro–motor coordination based on spatio–motor representations was investigated. It was found that piano majors effectively employed an external motor encoding strategy, in which auditory information was encoded as finger movements analogous to playing the piano. The motor encoding strategy was stable or robust against interference and time decay ; that is, the effects of finger movements became stronger as melody length and retention interval increased, and this was especially true when interference stimuli were interpolated during the retention interval, because the more challenging the situation became, the greater the efforts of the subjects to reduce the latency and to make elaborative and rapid finger movements. Some subjects could employ this external spatio–motor representation (motor encoding strategy) effectively on their own, while others could not employ it independently of verbal rehearsal of note names.

In Experiment 8, it was found that, when pitches of melodies were encoded, the visuo–spatial representation (i.e., visual image of spatial configurations) for pitch information were more accurate in the compatible spatial directions related to the motor system used in playing the instrument (internal spatio–motor representations). Piano majors would have spatial images in which the right and upper directions are compatible with higher pitch.

For violin majors, the forward direction is compatible with higher pitch. For violoncello majors, the lower and backward directions are compatible with higher pitch. For vocal music majors, the right, lower and backward directions are compatible with higher pitch. For less well musically trained subjects, the upper direction is compatible with higher pitch.

The evidence that while encoding pitches subjects who are highly trained in playing the instrument precisely tracked their internal visuo-spatial representations in the compatible spatial directions related to the motor system used in playing the instrument was obtained in the study of eye movements (Experiment 9).

An attempt is made to interpret these results in relation to Paivio's dual-coding theory, the levels-of-processing theory, and Baddeley's working memory model.

As described in Chapter 1 and Experiment 6, Paivio (1969, 1971, 1978, 1986) proposed the dual-coding theory, the essence of which is that there are two basic independent but interconnected systems for the representation and processing of information. The verbal system deals with linguistic information and stores it in an appropriate verbal form, while the nonverbal system carries out image-based processing and representation. Within the two systems there are associative structures. Both systems have basic representational units that are linked to one another by referential connections : Logogens for the verbal system and Imagens for the nonverbal system. Both the Logogens and Imagens are further divided into sub-systems (i.e., visual, auditory, kinesthetic, gustatory, and olfactory) which process either verbal or nonverbal information in the different modalities (modality-specific).

Considering the findings in Experiments 1, 4, and 6, for example, the dual–coding by pitch rehearsal along with note names is consistent with Paivio's theory, because pitch rehearsal would be an auditory Imagen and note names would be a auditory Logogen. If the triple–coding by visualizing staff notation along with pitch rehearsal with note names is consistent with the Paivio's theory, staff notation must be a visual Imagen, pitch rehearsal must be an auditory Imagen, and note names must be a auditory Logogen. Furthermore, the dual–coding by visualizing melodic contour along with pitch rehearsal would be considered the operation of a single system in Paivio's theory, because the melodic contour would be a visual Imagen and pitch rehearsal would be an auditory Imagen, and both codes would be together in a nonverbal system.

The results regarding explicit motor representations found in Experiment 7 provide some evidence supporting important views proposed in the levels–of–processing theory. The greater the degree of challenge in terms of melody length, ISI duration and interfering stimuli, the greater the efforts of the subjects to reduce the latency and to make elaborative and rapid finger movements. This finding supports the view that cognitive (conscious) effort is used for encoding (e.g., Battig, 1979 ; Jacoby & Craik, 1979 ; Lockhart, Craik, & Jacoby, 1976). As shown in Figure 7–9, it was also found that the finger movements gradually became more appropriate, accurate and rapid during the retention interval. This finding supports the view of a *continuum* of rehearsal operations running from the minimal processing necessary to repeat a word continuously to various types of elaborative processing involving either further enrichment of one item or associative linkage of several items (Craik, 1979), and the view that the encoding process gradually spreads and becomes richer and more elaborative (e.g., Anderson & Reder, 1979 ; Craik & Tulving, 1975 ;

Lockhart, Craik, & Jacoby, 1976). Craik and Tulving (1975) explained in terms of verbal stimulus that the minimal core encoding can be elaborated by a context of further structural, phonemic, and semantic encodings. Thus, the more the subjects rehearsed elaborative tapping, the greater the recognition performance became (e.g., Rundus, 1971, 1977, 1980 ; Rundus & Atkinson, 1970 ; Rundus, Loftus, & Atkinson, 1970).

Comparison of the motor rehearsal rate with the verbal and visual rehearsal rates revealed that the former is somewhat faster than the verbal rate, followed by the visual rate. The average rate of finger movement obtained in Experiment 7 was about 150 to 200 msec per tap, when the subjects made an effort to tap many times. The fastest rate was about 100 msec per tap. The data are almost the same as the rate for typewriting (Steinberg, Monsell, Knoll, & Wright, 1980 ; see Experiment 7). The verbal rehearsal rate is usually about three to six letters per second (e.g., Landauer, 1962 ; Norman, 1976 ; Thomas, Hill, Carroll, & Garcia, 1970 ; Warren, 1969), or about 160 to 340 msec per letter. Fixation duration of about 250 msec is necessary to perceive visual information (Rayner, 1978)[(10)]. Vaughn (1983) cites several studies that conclude that fixations must be at least 200 msec in duration in order for perception to take place (of which 100 to 140 msec is required reaction time after the decision has been made when and where to move the eyes). In Experiment 4, the subjects reported that when the eight circles appeared successively with a constant duration of 500 msec at each point, the duration was experienced as appropriate to track the circle movement[(5)]. In Experiment 5, the fixation duration data revealed that when the subjects precisely tracked their internal visual representation for pitch, the duration was about 300 to 430 msec per fixation.

When subjects are motivated to attempt to encode pitch information, the greater the number of times and faster the code is rehearsed, the deeper and more elaborately the code is processed. Motor encoding with finger movements might be relatively deep level of processing requiring much encoding effort at the first stage (it becomes gradually automatically), followed by verbal encoding with note names and, finally, visual encoding with melodic contour, which is a relatively shallow level of processing requiring less encoding effort (see Experiment 1). Visual encoding with staff notation might be somewhat deep level of processing, because pitches are necessary to be identified as their note names. The findings obtained in Experiments 1,4 to 7 would suggest that for the subjects highly trained in playing the instrument, motor encoding would be more stable or robuster against interference and time decay, and more effective for encoding pitches than verbal or visual encoding. Furthermore, motor encoding strategy could be employed independently of visual representations, while it may not be employed, in some cases, independently of verbal codes.

As described in Chapter 1, Baddeley proposed the working memory model (see Baddeley, 1986, 1990), which is herein applied to the interpretation of the encoding strategies found in this study.

The working memory system consists of three components, the most important of which is the central executive. It has limited capacity, and is used when dealing with most cognitively demanding tasks. Unfortunately, however, there has been very little clarification of the role played by the central executive. The articulatory loop and the visuo-spatial sketch pad are subordinate systems.

The articulatory loop, in which verbal rehearsal occurs, consists of a passive phonological store and an articulatory control process. Originally a passive phonological store is directly concerned with speech perception. The neuropsychological evidence indicates that it is an auditory short-term store in which the processing is at a phonological level, and meaningful sounds, continuous speech and visually presented materials are not processed. An articulatory suppression task has little effect on the storage of phonological material. An articulatory control process is linked to speech production. An articulatory suppression task effectively prevents the use of "inner voice" or subvocal rehearsal.

Logie and Edworthy (1986) investigated mechanisms of memory for tone sequences. Subjects were presented with pairs of tone sequences and instructed to make recognition judgments (see Baddeley & Logie, 1992). The task was performed on its own (control condition), or concurrently with each of three tasks ; articulatory suppression, word–nonword homophone judgments, and visual matching. The disruptive effects of concurrent tasks on the recognition performance seemed to support the view that articulatory suppression and homophone judgments involve separate mechanisms, and that memory for tone sequences involves both mechanisms ; i.e., a phonological store and subvocal rehearsal.

Based on these properties of the phonological store and the articulatory control process, the short–term store of pitch information with acoustic level (auditory imagery), in which 'raw' pitch is mentally rehearsed, might be analogous in function to the phonological store, and the verbal rehearsal of note names (at accurate pitches) might be analogous to subvocal "singing" in the articulatory control process..

The visuo–spatial sketch pad specialized for spatial and/or visual coding was defined by Baddeley (1986) as "a system especially well adapted to the storage of spatial information, much as a pad of paper might be used by someone trying for example to work out a geometric puzzle." As found in this study, visualization of pitch information was based on visual representations of melodic contour or staff notation, or visuo–spatial representations which were closely related to external or internal spatio–motor representations. The evidence that subjects tracked precisely their visuo–spatial image while visualizing pitch information was obtained by analyzing eye movements. These results are accounted for well by the visuo-spatial sketch pad concept.

Coltheart (1972) proposed a model for perception of letters under conditions of brief exposure in which two coding processes, one a fast visual coding process, and the other a slower, but more permanent abstract (or name) coding process, act simultaneously on the stimulus, The visual code is developed at the rate of about 10 msec per letter (iconic memory) and has a maximum capacity of four letters. At the same time, a representation of the letter names is developed at the rate of 100 msec per letter. These two representations coexist for a time.

Sloboda (1976) investigated recognition of briefly exposed pitch notation in musicians and non-musicians. Notes were presented on a staff, and after an interval, the subjects were required to reproduce (draw) the notes they had seen. Sloboda concluded that both musicians and non-musicians encode the notes visually at brief duration (20 msec), but that a second, nonvisual, code (a naming code or a pitch or acoustic code) comes into operation in musicians when the exposure time is increased beyond 150 msec. These results are accounted for well by the model proposed by Coltheart (1972).

The purpose of the present study was to investigate the encoding strategies for auditory pitch information, therefore, pitch was presented auditorily instead of visual notations. Therefore, there is a possibility that the first code is acoustic pitch and that a naming, visual or motor code then comes into operation. This study revealed the existence of multi-coding, but could not reveal the order in which several codes come into operation. If the order or duration in which each code is operating were found, the mechanism by which auditory information is translated into codes would become clear.

As described in Chapter 1, Posner (1973) proposed concerning codes in memory that (a) there are at least three types of codes : visual, verbal, and motor (Bower, 1972b) ; (b) each code endures, and is not a transient residual of stimulation ; (c) people differ in their propensity to use each type of code ; (d) these codes are parts of separate memory systems that can be examined in isolation in the laboratory. In this study, all four points were experimentally demonstrated concerning encoding of pitch information in short melodies. Music has a hierarchical structure which consists of several levels similar to language (Chapter 1). Experiments in the laboratory on short-term memory for short melodies correspond to the second dimension of the hierarchical structure proposed by Umemoto (1990). Studies on this dimension might lay the foundations of an understanding of memory for music in everyday situations.

****

(10) In fluent reading the duration of each fixation is about 250 msec, and the eye moves from one fixation to another in a rapid sweep known as a saccade, lasting about 50 msec. The available evidence suggests that we take in visual information only during the static fixations and not during saccadic movements (Rayner, 1978).

REFERENCES

Abe, J. 1987 Perception and prediction in melody : Simulation of the final–tone extrapolating behavior. *The Japanese Journal of Psychology*, **58**, 275–281.

Abe, J. & Hoshino, E. 1985 Schema–driven properties in melody cognition : Experiments on final–tone extrapolation by music experts. *The Japanese Journal of Psychonomic Science*, **4**, 1–9.

Aiello, R. 1994 Music and language : Parallels and contrasts. In R. Aiello & J.A. Sloboda (Eds.), *Musical Perceptions*. Oxford University Press. Pp. 40–63.

Anderson, J.R. & Reder, L.M. 1979 An elaborative processing explanation of depth of processing. In L.S. Cermak & F.I.M. Craik (Eds.), *Levels of processing in human memory*. Lawrence Erlbaum Associates. Pp. 385–404.

Atkinson, R.C. & Shiffrin, R.M. 1968 Human memory : A proposed system and its control processes. In K.W. Spence & J.T. Spence (Eds.), *The psychology of learning and motivation* **2**. London : Academic Press.

Atkinson, R.C. & Shiffrin, R.M. 1971 The control of short–term memory. *Scientific American*, **225**, 82–90.

Baddeley, A.D. 1978 The trouble with levels : A reexamination of Craik and Lockhart's framework for memory research. *Psychological Review*, **85**, 139–152.

Baddeley, A.D. 1986 *Working Memory*. Oxford : Oxford University Press.

Baddeley, A.D. 1990 *Human Memory : Theory and practice*. Hove and London : Lawrence Erlbaum Associates.

Baddeley, A.D. & Hitch, G. 1974 Working memory. In G.H. Bower (Ed.), *The Psychology of learning and motivation* **8**. London : Academic Press.

Baddeley, A.D. & Lieberman, K. 1980 Spatial working memory. In R. Nickerson (Ed.), *Attention and Performance* **8**. Hillsdale, N.J. : Lawrence Erlbaum Associates.

Baddeley, A.D. & Logie, R.H. 1992 Auditory imagery and working memory. In D. Resberg (Ed.), *Auditory Imagery*. Hillsdale, N.J. : Lawrence Erlbaum Associates. Pp. 179–197.

Baily, J. 1985 Music structure and human movement. In P. Howell, & I. Cross, & R. West (Eds.), *Musical Structure and Cognition*. Academic Press. Pp. 237–258.

Baker, M.A. & Loeb, M. 1973 Implications of measurements of eye fixations for a psychophysics of form perception. *Perception and Psychophysics*, **13**, 185–192.

Bartlett, J.C. & Dowling, W.J. 1980 The recognition of transposed melodies : A key-distance effect in developmental perspective. *Journal of Experimental Psychology : Human Perception and Performance*, **6**, 501–515.

Battig, W.F. 1979 The flexibility of human memory. In L.S. Cermak & F.I.M. Craik (Eds.), *Level of processing in human memory*. New York : Lawrence Erlbaum Associates. Pp. 23–44.

Bellugi, U., Klima, E.S., & Siple, P.A. 1975 Remembering in signs. *Cognition*, **3**, 93–125.

Bennett, A., Hausfeld, R.A., Hausfeld, S., Reeve, R.A., & Smith, R.J. 1978 Recognizing musical patterns. *Work shops in cognitive processes*. Curtis Brown Ltd. Pp. 97–101.

Bever, T.G., & Chiarello, R.J. 1974 Cerebral dominance in musicians and non-musicians. *Science*, **185**, 537–539.

Bharucha, J.J. 1987 Music cognition and perceptual facilitation : A connectionist framework. *Music Perception*, **5**, 1–30.

Bharucha, J.J. & Stoeckig, K. 1986 Reaction time and musical expectancy : Priming of chords. *Journal of Experimental Psychology : Human Perception and Performance*, **12**, 403–410.

Bharucha, J.J. & Stoeckig, K. 1987 Priming of chords : Spreading activation or overlapping frequency specta ? *Perception and Psychophysics*, **41**, 519–524.

Bogen, J.E. & Gordon, H.W. 1971 Musical tests for functional lateralization with intracarotid amobarbital, *Nature*, **230**, 524–525.

Boltz, M. 1989 Perceiving the end : effects of tonal relationships on melodic completion. *Journal of Experimental Psychology : Human Perception and Performance*, **15**, 749–761.

Bower, G.H. 1967 A multicomponent theory of the memory trace. In K.W. Spence & J.T. Spence (Eds.), *The psychology of learning and motivation* **1**. New York : Academic Press.

Bower, G.H. 1970 Analysis of a mnemonic device. *American Scientist*, **58**, 496–510.

Bower, G.H. 1972a Mental imagery and associative learning. In L. Gregg (Ed.), *Cognition in learning and memory*. New York : Wiley.

Bower, G.H. 1972b Stimulus–sampling theory of encoding variability. In A.W. Melton & E. Martin (Eds.), *Coding processes in human memory*. Washington,D.C.: Winston & Sons.

Broadbent, D.E. 1954 The role of auditory localization in attention and memory span. *Journal of Experimental Psychology*, **47**, 191–196.

Brooks, L.R. 1968 Spatial and verbal components of the act of recall. *Canadian Journal of Psychology*, **22**, 349–368.

Brown, M. 1979 *Left handed : Right handed*. David & Charles Ltd., London.

Burns, E.M. & Ward, W.D. 1974 Categorical perception of musical intervals. *Journal of the Acoustical Society of America*, **55**, 456(A).

Burns, E.M. & Ward, W.D. 1978 Categorical perception —— phenomenon or epiphenomenon : Evidence from experiments in the perception of melodic musical intervals. *Journal of the Acoustical Society of America*, **63**, 456–468.

Burns, E.M. & Ward, W.D. 1982 Intervals, Scales, and Tuning. In D. Deutsch (Ed.), *The Psychology of Music*. Academic Press. Pp. 241–269.

Carpenter, P.A. & Just, M.A. 1977 Reading comprehension as the eye sees it. In M.A. Lust & P.A. Capenter (Eds.), *Cognitive processes in comprehension*. Hillsdale,N.J. : Erlbaum. Pp. 109–139.

Cohen, G. 1973 Hemispheric differences in serial versus parallel processing. *Journal of Experimental Psychology*, **97**, 349–356.

Cohen, G. 1977 *The psychology of cognition*. London & New York : Academic Press.

Cohen, R.L. 1981 On the generality of some memory laws. *Scandinavian Journal of Psychology*, **22**, 267–281.

Cohen, R.L. 1983 The effect of encoding variables on the free recall of words and action events. *Memory and Cognition,* **11**, 575–582.

Coltheart, M. 1972 Visual information processing. In B.M. Foss (Ed.), *New Horizons in Psychology* **2**. London : Penguin.

Conrad, R. 1964 Acoustic confusions in immediate memory. *British Journal of Psychology*, **55**, 75–84.

Conrad, R. 1972 Short–term memory in the deaf : A test for speech coding. *British Journal of Psychology*, **63**, 173–180.

Cooper, L.A. & Shepard, R.N. 1973 Chronometric studies of the rotation of mental images. In W.G. Chase (Ed.), *Visual Information Processing*. New York : Academic Press.

Corballis, M.C. & Beale, I.L. 1976 *The Psychology of Left and Right*. Lawrence Erlbaum Associates.

Corman, C.N. & Wickens, D.D. 1968 Retroactive inhibition in short-term memory. *Journal of Verbal Learning and Verbal Behavior*, **7**, 16–19.

Craik, F.I.M. 1979 Human Memory. *Annual Review of Psychology*, **30**, 69–102.

Craik, F.I.M. & Lockhart, R.S. 1972 Levels of processing : A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, **11**, 671–684.

Craik, F.I.M. & Tulving, E. 1975 Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology : General*, **104**, 268–294.

Craik, F.I.M. & Watkins, M.J. 1973 The role of rehearsal in short-term memory. *Journal of Verbal Learning and Verbal Behavior*, **12**, 599–607.

Cuddy, L.L. & Cohen, A.J. 1976 Recognition of transposed melodic sequences. *Quarterly Journal of Experimental Psychology*, **28**, 255–270.

Cuddy, L.L., Cohen, A.J., & Miller, J. 1979 Melody recognition : The experimental application of musical rules. *Canadian Journal of Psychology*, **33**, 148–157.

Damasio, A.R., Almeida, L., & Damasio, H. 1975 Nervous function after right hemispherectomy, *Neurology*, **25**, 89–93

Darwin, C.J., Turvey, M.T., & Crowder, R.G. 1972 An auditory analogue of the Sperling partial report procedure : Evidence for brief auditory storage. *Cognitive Psychology*, **3**, 255–267.

Davies, J.B. & Jennings, J. 1977 Reproduction of familiar melodies and the perception of tonal sequences. *Journal of the Acoustical Society of America*, **61**, 534–541.

Davison, L. & Scripp, L. 1988 Young children's musical representations : windows on music cognition. In J.A. Sloboda (Ed.), *Generative Processes in Music*. Oxford University Press. Pp. 195–230.

Deatherage, B.H. & Evans, T.R. 1969 Binaural masking : backward, forward and simultaneous effects. *Journal of the Acoustical Society of America*, **46**, 362–371.

Deutsch, D. 1970 Tones and numbers : Specificity of interference in short-term memory. *Science*, **168**, 1604–1605.

Deutsch, D. 1972 Mapping of interactions in the pitch memory store. *Science*, **175**, 1020–1022.

Deutsch, D. 1977 Memory and attention in music. In M. Critchley & R.A. Henson (Eds.), *Music and the Brain*. London : Heinemann, Pp. 95–130.

Deutsch, D. 1980 The processing of structured and unstructured tonal sequences. *Perception and Psychophysics*, **28**, 381–389.

Deutsch, D. 1981 The internal representation of pitch sequences in tonal music. *Psychological Review*, **88**, 503–522.

Deutsch, D. 1982a Introduction. In D. Deutsch (Ed.), *The Psychology of Music*. Academic Press.

Deutsch, D. 1982b Grouping mechanisms in music. In D. Deutsch (Ed.), *The Psychology of Music*. Academic Press. Pp. 99–134.

Deutsch, D. 1982c The processing of pitch combinations. In D. Deutsch (Ed.), *The Psychology of Music*. Academic Press. Pp. 271–316.

Deutsch, D. & Feroe, J. 1981 The internal representation of pitch sequences in tonal music. *Psychological Review*, **88**, 503–522.

Dewar, K.M., Cuddy, L.L., & Mewhort, D.J.K. 1977 Recognition memory for single tones with and without context. *Journal of Experimental Psychology : Human Learning and Memory*, **3**, 60–67.

Dewitt, L.A. & Crowder, R.G. 1986 Recognition of novel melodies after brief Delays. *Music Perception*, **3**, 259–274.

Dowling, W.J. 1971 Recognition of inversions of melodies and melodic contours. *Perception and Psychophysics*, **9**, 348–349.

Dowling, W.J. 1972 Recognition of melodic transformations : Inversion, retrograde, and retrograde–inversion. *Perception and Psychophysics*, **12**, 417–421.

Dowling, W.J. 1978 Scale and contour : Two components of a theory of memory for melodies. *Psychological Review*, **85**, 341–354.

Dowling, W.J. 1982 Melodic information processing and its development. In D. Deutsch (Ed.), *The Psychology of Music*. Academic Press. Pp. 413–429.

Dowling, W.J. 1988 Tonal structure and children's early learning of music. In J.A. Sloboda (Ed.), *Generative Processes in Music*. Oxford University Press. Pp. 113–128.

Dowling, W.J. 1991 Pitch structure. In P. Howell, R. West, & I. Cross (Eds.), *Representing musical structure*. Academic Press. Pp. 33–57.

Dowling, W.J. 1994 Melodic contour in hearing and remembering melodies. In R. Aiello & J.A. Sloboda (Eds.), *Musical Perceptions*. Oxford University Press. Pp. 173–190.

Dowling, W.J. & Bartlett, J.C. 1981 The importance of interval information in long–term memory for melodies. *Psychomusicology*, **1**, 30–49.

Dowling, W.J. & Fujitani, D.A. 1971 Contour, interval, and pitch recognition in memory for melodies. *Perception and Psychophysics*, **9**, 524–531.

Drake, R.M. 1954 *Drake Musical Aptitude Tests.* Chicago : Science Research Associate.

Edworthy, J. 1985 Melodic contour and musical structure. In P. Howell, & I. Cross, & R. West (Eds.), *Musical Structure and Cognition.* Academic Press. Pp. 169–188.

Efron, R. 1970 Effect of stimulus duration on perceptual onset and offset latencies. *Perception and Psychophysics*, **8**, 231–234.

Engelkamp, J. 1986 Nouns and verbs in paired-associate learning : Instructional effects. *Psychological Research*, **48**, 153–159.

Engelkamp, J. 1988 Modality-specific encoding and class in verbal learning. In M.M. Gruneberg, P.E. Morris, & R.N. Sykes (Eds.), *Practical aspects of memory : current issues*, Vol. **1**, Chichester : John Wiley. Pp. 415–420.

Engelkamp, J. 1990 Memory of action events : Some implications for memory theory and for imagery. In C. Cornoldi & M.McDaniel (Eds.), *Imagery and Cognition.* New York, Berlin, Heidelberg : Springer. Pp. 415–420.

Erickson, R. 1984 A perceptual structure for tonal centering. *Music Perception,* **2**, 1–5.

Eysenck, M.W. 1978 Verbal remembering. In B.M. Foss (Ed.), *Psychology survey* **1**. London : Allen & Unwin.

Eysenck, M.W. 1979 Depth, elaboration, and distinctiveness. In L.S. Cermak & F.I.M. Craik (Eds.), *Levels of processing in human memory.* Lawrence Erlbaum Associates.

Gardner, M. 1964 *The Ambidextrous Universe : Left, Right, and the Fall of Parity.* Basic Books, Inc.

Gates, A. & Bradshaw, J.L. 1977a Music perception and cerebral asymmetries. *Cortex,* **13**, 390–401.

Gates, A. & Bradshaw, J.L. 1977b The role of the cerebral hemispheres in music. *Brain and Language*, **4**, 403–431.

Gates, A., Bradshaw, J.L., & Nettleton, N.C. 1974 Effect of different delayed auditory feedback intervals on a music performance task. *Perception and Psychophysics*, **15**, 21–25.

Gazzaniga, M.S. & LeDoux, J.E. 1978 *The Integrated Mind*. Plenum Press. New York.

Geschwind, N. & Levitsky, W. 1968 Human brain : Left–right asymmetries in temporal speech region, *Science*, **161**, 186–187.

Goolsby, T. 1989 Computer applications to eye movement research in music reading. *Psychomusicology*, **8**, 111–126.

Gordon, E. 1965 *Musical Aptitude Profile*. New York : Houghton Mifflin Co.

Gordon, H.W. 1970 Hemispheric asymmetries in the perception of musical chords. *Cortex*, **6**, 387–398.

Gordon, H.W. 1978 Hemispheric asymmetry for dichotically presented chords in musicians and non–musicians, males and females. *Acta Psychologica*, **42**, 383–395.

Gordon, H.W. 1980 Degree of Ear Asymmetries for perception of dichotic chords and for illusory chord localization in musicians of different levels of competence. *Journal of Experimental Psychology : Human Perception and Performance*, **6**, 516–527.

Gordon, H.W. 1983 Music and the right hemisphere. In A.W. Young (Ed.), *Function of the right cerebral hemisphere*. Academic Press. Pp. 65–86.

Gourevitch, V. & Galanter, E. 1967 A significance test for one parameter isosensitivity functions. *Psychometrika*, **32**, 25–33.

Hatta, T. 1982 *Study of cerebral hemispheric differences*. Kazama Shobo.

Hatta, T. & Nakatsuka, Z. 1975 Hatta–Nakatsuka Handedness Inventory. In S. Ono (Ed.), *Papers in cerebration of 63rd birthday of Prof. K.Ohnishi.* Osaka City University, Pp. 224–247.

Hebb, D.O. 1968 Concerning imagery. *Psychological Review,* **75**, 466–477.

Hoshino, E. 1985 Influence of melodic "context" on recognition memory pitch. *The Japanese Journal of Psychology,* **56**, 132–137.

Hoshino, E. & Abe, J. 1981 Tonality and the coherence in melody cognition. *Hokkaido Behavioral Science Report,* Series P (Supplement), **23**, 1–19.

Hoshino, E. & Abe, J. 1984 "Tonality" and the final tone extrapolation in melody cognition. *The Japanese Journal of Psychology,* **54**, 344–350.

Hubbard, T.L. & Stoeckig, K. 1992 The representation of pitch in musical images. In D. Reisberg (Ed.), *Auditory Imagery.* Hillsdales, N.J. : Lawrence Erlbaum Associates. Pp. 199–225.

Idson, W.L. & Massaro, D.W. 1978 A bidimensional model of pitch in the recognition of melodies. *Perception and Psychophysics,* **24**, 551–565.

Jacoby, L.L. & Craik, F.I.M. 1979 Effects of elaboration of processing at encoding and retrieval : Trace distinctiveness and recovery of initial context. In L.S. Cermak & F.I.M. Craik (Eds.), *Level of processing in human memory.* New York : Lawrence Erlbaum Associates. Pp. 1–20.

Jenkins, J.J. 1974 Remember that old theory of memory ? Well forget it !, *American Psychologist*, **29**, 785–795.

Jones, M.R. 1976 Levels of structure in the reconstruction of temporal and spatial serial patterns. *Journal of Experimental Psychology : Human Learning and Memory,* **2**, 475–488.

Kallman, H.J. & Massaro, D.W. 1979 Tone chroma is functional in melody recognition. *Perception and Psychophysics,* **26**, 32–36.

Keppel, G. & Underwood, B.J. 1962 Proactive inhibition in short-term retention of single items. *Journal of Verbal Learning and Verbal Behavior*, **1**, 153-161.

Kimura, D. 1961 Some effects of temporal lobe damage on auditory perception. *Canadian Journal of Psychology*, **15**, 156-165.

Kimura, D. 1963 Speech lateralization in young children as determined by an auditory test. *Journal of Comparative and Physiological Psychology*, **56**, 899-902.

Kimura, D. 1964 Left-right differences in the perception of melodies. *Quarterly Journal of Experimental Psychology*, **16**, 355-358.

Kimura, D. 1967 Functional asymmetry of the brain in dichotic listening. *Cortex*, **3**, 163-178.

Kimura, D. 1968 Neural processing of backwards-speech sounds, *Science*, **161**, 395-396.

Kimura, D. 1973 The asymmetry of the human brain, *Scientific American*, **228**, 70-78.

Kintsh, W. & Buschke, H. 1969 Homophones and synonyms in short-term memory. *Journal of Experimental Psychology*, **80**, 403-407.

Kosslyn, S.M. 1973 Scanning visual image : Some structural implications. *Perception and Psychophysics*, **14**, 90-95.

Kosslyn, S.M. 1975 Information representation in visual image. *Cognitive Psychology*, **7**, 341-370.

Kosslyn, S.M. 1980 *Image and mind*. Cambridge, Mass. : Harvard University Press.

Kroll, N.E.A. 1972 Short-term memory and the nature of interference from concurrent shadowing. *Quarterly Journal of Experimental Psychology*, **24**, 414-419.

Kroll, N.E.A. 1975 Visual short-term memory. In D. Deutsch & J.A. Deutsch (Eds.), *Short-term memory*. Academic Press.

Kroll, N.E.A. & Kellicutt, M.H. 1972 Short-term recall as a function of covert rehearsal and of intervening task. *Journal of Verbal Learning and Verbal Behavior*, **11**, 196-204.

Kroll, N.E.A., Parks, T., Parkinson, S.R., Bieber, S.L., & Johnson, A.L. 1970 Short-term memory while shadowing : Recall of visually and of aurally presented letters. *Journal of Experimental Psychology*, **85**, 220-224.

Krumhansl, C.L. 1979 The psychological representation of musical pitch in a tonal context. *Cognitive Psychology*, **11**, 346-374.

Krumhansl, C.L. 1990 *Cognitive foundations of musical pitch*. Oxford Psychology Series, **17**. Oxford University Press.

Krumhansl, C.L. 1991 Melodic structure : Theoretical and empirical descriptions. In J. Sundberg, L. Nord, & R. Carlson (Eds.), *Music, Language, Speech and Brain*. Macmillan press. Pp. 269-283.

Krumhansl, C.L. & Kessler, E.J. 1982 Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review*, **89**, 334-368.

Krumhansl, C.L. & Shepard, R. 1979 Quantification of the hierarchy of tonal functions within a diatonic context. *Journal of Experimental Psychology : Human Perception and Performance*, **5**, 579-594.

LaBerge, D. 1975 Acquisition of automatic processing in perceptual and associative learning. In P.M.E. Rabbit & S. Dronic (Eds.), *Attention and Performance* **5**. New York : Academic Press. Pp. 55-64.

Landauer, T.K. 1962 Rate of implicit speech. *Perception and Motor Skills*, **15**, 646.

Lebrun, Y. & Zangwill, O. 1981 *Lateralisation of Language : In the child.* Swets & Zeitlinger B,V., Lisse.

Lerdahl, F. 1988 Tonal pitch space. *Music Perception*, **5**, 315-350.

Lerdahl, F. & Jackendoff, R. 1983 *A Generative Theory of Tonal Music.* Cambrige, M.A. : MIT. Press.

Levin, J.R. 1976 What have we learned about maximizing what children learn ? In J.R. Levin & V.L. Allen (Eds.), *Cognitive learning in children.* New York : Academic Press.

Levy, V.A. 1971 Role of articulation in auditory and visual short-term memory. *Journal of Verbal Learning and Verbal Behavior*, **10**, 123-132.

Lindsay, P.H. & Norman, D.A. 1977 *Human Information Processing : An Introduction to Psychology.* (2nd Edition). Academic Press, Inc. New York.

Locke, J.L. & Locke, V.L. 1971 Deaf children's phonetic, visual, and dactylic coding in a grapheme recall task. *Journal of Experimental Psychology*, **89**, 142-146.

Locke, S. & Kellar, L. 1973 Categorical perception in a non-linguistic mode. *Cortex*, **9**, 355-368.

Lockhart, R.S., Craik, F.I.M., & Jacoby, L. 1976 Depth of processing ; recognition and recall. In J.Brown (Ed.), *Recall and recognition.* Wiley. Pp. 75-102.

Loftus, G.R. 1976 A framework for a theory of picture recognition. *Eye movements and psychological processes.* R.A. Monty & J.W. Senders (Eds.), Hillsdale,N.J. : Lawrence Erlbaum Associates.

Logie, R.H. & Edworthy, J. 1986 Sheared mechanisms in the processing of verbal and musical material. In D.G. Russell, D. Marks, & J. Richardson (Eds.), *Imagery* **2.** Dunedin. New Zealand : Human Performance Associates. Pp. 33-37.

Longuet-Higgins, H.C. 1978 The perception of music. *Interdisciplinary Science Review*, **3**, 148-156.

Macmillan, N.A. & Creelman, C.D. 1991 *Detection Theory : A User's Guide.* Cambridge University Press.

Marin, O.S. 1982 Neurological aspects of music perception and performance. In D. Deutsch (Ed.), *The Psychology of Music*. Academic Press.

Marks, D.F. 1973 Visual imagery differences and eye movements in the recall of pictures. *Perception and Psychophysics*, **14**, 407–412.

Massaro, D.W. 1970 Retroactive interference in short–term recognition memory for pitch. *Journal of Experimental Psychology*, **83**, 32–39.

Massaro, D.W. 1972 Perceptual images, processing time, and perceptual units in auditory perception. *Psychological Review*, **79**, 124–145.

McConkie, G.W. 1983 Eye movements and perception during reading. In K. Rayner (Ed.), *Eye movements in reading : Perceptual and language processes*. New York : Academic Press. Pp. 65–96.

Mckee, G., Humphrey, B., & McAdam, D.W. 1973 Scaled lateralization of alpha activity during linguistic and musical tasks. *Psychophysiology*, **10**, 441–443.

Mikumo, M. 1987a Effects of training in music on processing melodies, —— Processing mode of recognition tasks ——. *Nara Women's University, Unpublished Master's Thesis in Psychology.*

Mikumo, M. 1987b Cerebral dominance in processing of music. *Proceedings of the 99th Convention of Kansai Psychological Association*, 13.

Mikumo, M. 1988 Effects of training in music for processing of tonal and atonal melodies. *Konan Women's University Laboratory of Psychology Annual Reports*, **6**, 65–80.

Mikumo, M. 1989a Coding strategies of tonal and atonal melodies. *Proceedings of the first International Conference on Music Perception and Cognition*. Kyoto, JAPAN. 147–150.

Mikumo, M. 1989b Coding strategies and recognition of melodies. *Proceedings of the 53rd Congress of Japanese Psychological Association*, 687.

Mikumo, M. 1990a Construction of a test for cognition of pitch deviation. —— Analysis of pitch interval and factor analysis ——. *Proceedings of the 54th Congress of Japanese Psychological Association*, 567.

Mikumo, M. 1990b Coding strategies and recognition of melodies. *The Japanese Journal of Psychology*, **61**, 291-298.

Mikumo, M. 1991a Encoding strategy of tapping for pitches of melodies (No.1). *Proceedings of the 55th Congress of Japanese Psychological Association*, 215.

Mikumo, M. 1991b Effects of training in music on hemispheric functional differences for processing of melodies and chords. *Human Developmental Research*, **7**, 167-179.

Mikumo, M. 1992a Finger-tapping strategy for pitch encoding of melodies. *Abstract of the 2nd International Conference on Music Perception and Cognition*. UCLA, U.S.A. 63.

Mikumo, M. 1992b Finger-tapping Strategy for encoding melodies. *Proceedings of International Symposium on Musical Acoustics*. Tokyo, JAPAN. 195-198.

Mikumo, M. 1992c Encoding strategy of tapping for pitches of melodies (No.2). *Proceedings of the 56th Congress of Japanese Psychological Association*, 782.

Mikumo, M. 1992d Encoding strategies for tonal and atonal melodies. *Music Perception*, **10**, 73-81.

Mikumo, M. 1993a Visual encoding strategy for pitches of melodies. *Proceedings of the 57th Congress of Japanese Psychological Association*, 301.

Mikumo, M. 1993b Dichotic listening method (Report of Workshop). *The 57th Congress of Japanese Psychological Association.*

Mikumo, M. 1994a Image of spatial configuration for pitch encoding. *Proceedings of the third International Conference on Music and Cognition.* Liege. Belgium. 447-448.

Mikumo, M. 1994b Finger-tapping for pitch encoding of melodies. *Japanese Psychological Research*, **36**, 53-64.

Mikumo, M. 1994c Motor encoding strategy for pitches of melodies. *Music Perception*, **12**, 175-197.

Mikumo, M. 1994d Spatial image for pitch encoding. *Proceedings of the 58th congress of Japanese Psychological Association*, 725.

Mikumo, M. & Umemoto, T. 1988 Effects of training in music for retrieval of tone series. *Proceedings of the 52nd Congress of Japanese Psychological Association*, 624

Mikumo, M. & Umemoto, T. 1990 Construction of a test for cognition of pitch deviation. *Report of the Japanese Acoustical Society*, (H.90-2, MA-89-20).

Milner, B. 1962 Laterality effects in audition. In V.B. Mountcastle (Ed.), *Interhemispheric relations and cerebral dominance*. Baltimore : Johns Hopkins University Press. Pp. 177-195.

Milner, B. 1968 Lateralized suppression of dichotically presented digits after commissural section in man, *Science*, **161**, 184-186.

Morton, J. 1969 Interaction of information in word recognition. *Psychological Review*, **76**, 165-178.

Muter. P. 1980 Very rapid forgetting. *Memory and Cognition*, **8**, 174-179.

Neisser, U. 1967 *Cognitive psychology*. New York : Appleton-Century-Crofts.

Norman, D.A. 1976 *Memory and attention*. 2nd edition. John Wiley.

Norman, D.A. & Bobrow, D.G. 1975 On data-limited and resource-limited processes. *Cognitive Psychology*, **7**, 44-64.

Norman, D.C. 1986 *The Brian Code : Mechanisms of Information. Transfer and the Role of the Corpus Callosum*. Kinokuniya.

Ohgishi, M. 1978 Hemispheric differences in the mode of information processing. *The Japanese Journal of Psychology*, **49**, 257–264.

Ohta, N. 1993 Study on action–event memory (1). *Proceedings of the 57th Congress Japanese Psychological Association*, 762.

Paivio, A. 1969 Mental imagery in associative learning and memory. *Psychological Review*, **76**, 241–263.

Paivio, A. 1971 *Imagery and verbal processes*. New York : Holt, Rinehart & Winston.

Paivio, A. 1978 A dual coding approach to perception and cognition. In H.L. Pick & E. Saltzman (Eds.), *Modes of perceiving and processing information*. John Wiley. Pp. 39–51.

Paivio, A. 1986 *Mental representations : a dual coding approach*. Oxford University Press.

Parkinson, S.R. 1972 Short–term memory while shadowing : Multiple–item recall of visually and of aurally presented letters. *Journal of Experimental Psychology*, **92**, 256–265.

Parkinson, S.R., Parks, T.E., & Kroll, N.E.A. 1971 Visual and auditory short–term memory : Effects of phonemically similar auditory shadow material during the retention interval. *Journal of Experimental Psychology*, **87**, 274–280.

Peterson, L.R. & Johnson, S.T. 1971 Some effects of minimizing articulation on short–term retention. *Journal of Verbal Learning and Verbal Behavior*, **10**, 346–354.

Posner, M.I. 1973 Coordination of internal codes. In W.G. Chase (Ed.), *Visual information processing*. New York : Academic Press.

Pratt, C.C. 1929 The spatial character of high and low tones. *Journal of Experimental Psychology*, **13**, 278–285.

Proctor, R.W. & Fagnani, S.A. 1978 Effects of distractor stimulus modality in the Brown–Peterson distractor task. *Journal of Experimental Psychology : Human Learning and Memory*, **4**, 676–684.

Pylyshyn, Z. 1973 What the mind's eye tells the mind's brain. *Psychological Bulletin*, **80**, 1–24.

Rakowski, A. 1976 Tuning of isolated musical intervals. *Journal of the Acoustical Society of America*, **59**, S50(A).

Rayner, K. 1978 Foveal and parafoveal cues in reading. In J. Requin (Ed.), *Attention and Performance* **7**. Erlbaum, Hillsdale, New Jersey.

Reitman, J.S. 1971 Mechanisms of forgetting in short–term memory. *Cognitive Psychology*, **2**, 185–195.

Reitman, J.S. 1974 Without surreptitious rehearsal, information in short–term memory decays. *Journal of Verbal Learning and Verbal Behavior*, **13**, 365–377.

Repp, B.H. 1991 Some cognitive and perceptual aspects of speech and music. In J. Sundbderg, L. Nord, & R.Carlson (Eds.), *Music, Language, Speech and Brain*, Macmillan press. Pp. 257–268.

Rumelhart, D.E. & Norman, D.A. 1982 Simulating a skilled typist : a study of skilled cognitive–motor performance. *Cognitive Science*, **6**, 1–36.

Rundus, D. 1971 Analysis of rehearsal processes in free recall. *Journal of Experimental Psychology*, **89**, 63–77.

Rundus, D. 1977 Maintenance rehearsal and single–level processing. *Journal of Verbal Learning and Verbal Behavior*, **16**, 665–681.

Rundus, D. 1980 Maintenance rehearsal and long–term recency. *Memory and Cognition*, **8**, 226–230.

Rundus, D. & Atkinson, R.C. 1970 Rehearsal processes in free recall, a procedure for direct observation. *Journal of Verbal Learning and Verbal Behavior*, **9**, 99–105.

Rundus, D., Loftus, G.R., & Atkinson, R.C. 1970 Immediate free recall and three-week delayed recognition. *Journal of Verbal Learning and Verbal Behavior*, **9**, 684–688.

Saltz, E. & Dixon, D. 1982 Let's pretend : the role of motoric imagery in memory for sentences and words. *Journal of Experimental Child Psychology*, **34**, 77–92.

Saltz, E. & Donnenwerth–Nolan, S.I. 1981 Does motoric imagery facilitate memory for sentences ? A selective interference test. *Journal of Verbal Learning and Verbal Behavior*, **20**, 322–332.

Salzberg, P.M., Parks, T.E., Kroll, N.E.A., & Parkinson, S.R. 1971 Retroactive effects of phonemic similarity on short–term recall of visual and auditory stimuli. *Journal of Experimental Psychology*, **91**, 43–46.

Sasaki, M. & Watanabe, A. 1983 An experimental study of spontaneous writinglike behavior ("kusho") in Japanese. *The Japanese Journal of Educational Psychology*, **31**, 1–10.

Scheid, P. & Eccles, J.C. 1973 Music and speech : artistic functions of the human brain. *Psychology of Music*, **1**, 21–35.

Seamon, J.G. 1973 Coding and retrieval processes and the hemispheres of the brain. In S.J. Dimond & J.G. Beaumont (Eds.), *Hemisphere function in the Human Brain*. London : Paul Elek Ltd.

Seamon, J.G. & Gazzaniga, M.S. 1973 Coding strategies and cerebral laterality effects. *Cognitive Psychology*, **5**, 249–256.

Seashore, C.E. 1919 *Seashore Measures of Musical Talent*. Chicago : Stoelting.

Seashore, C.E., Lewis, D., & Saetveit, J.C. 1960 *Manual of instructions and interpretations for the Seashore Measures of Musical Talents* (2nd Revision). New York: The Psychological Corporation.

Sedgwick, H.A. & Festinger, L. 1976 Eye movements, efference, and visual perception. *Eye movements and psychological processes*. R.A. Monty & J.W. Senders (Eds.), Hillsdale, N.J. : Lawrence Erlbaum Associates.

Shaffer, L.H. 1973 Latency mechanisms in transcription. In S.Kornblum (Ed.), *Attention and Performance* **4**. New York : Academic Press. Pp. 435–446.

Shaffer, L.H. 1975 Control processes in typing. *Quarterly Journal of Experimental Psychology*, **27**, 419–432.

Shaffer, L.H. & Hardwick, J. 1970 The basis of transcription skill. *Journal of Experimental Psychology*, **84**, 424–440.

Shankweiler, D. 1966 Effects of temporal-lobe damage on perception of dichotically presented melodies. *Journal of Comparative and Physiological Psychology*, **62**, 115–119.

Shepard, R.N. 1982a Structure representations of musical pitch. In D. Deutsch (Ed.), *The Psychology of Music*. Academic Press. Pp. 344–390.

Shepard, R.N. 1982b Geometrical approximations to the structure of musical pitch. *Psychological Review*, **89**, 305–333.

Shepard, R.N. & Metzler, J. 1971 Mental rotation of three-dimensional objects. *Science*, **171**, 701–703.

Sherman, R. & Knight, M. 1977 *Music Theory Placement Test*. Muncie: Music School, Ball State University.

Shiffrin, R.M. 1973 Information persistence in short-term memory. *Journal of Experimental Psychology*, **100**, 39–49.

Shulman, H.G. 1970 Encoding and retention of semantic and phonemic information in short-term memory. *Journal of Verbal Learning and Verbal Behavior*, **9**, 499-508.

Shulman,H.G. 1972 Semantic confusion errors in short-term memory. *Journal of Verbal Learning and Verbal Behavior*, **11**, 221-227.

Shuter-Dyson, R. 1982 Musical ability. In D.Deutsch (Ed.), *The Psychology of Music*. Academic Press.

Siegel, J.A. 1974 Sensory and verbal coding strategies in subjects with absolute pitch. *Journal of Experimental Psychology*, **103**, 37-44.

Siegel, J.A. & Siegel, W. 1977a Absolute identification of notes and intervals by musicians. *Perception and Psychophysics*, **21**, 143-152.

Siegel, J.A. & Siegel, W. 1977b Categorical perception of tonal intervals. *Perception and Psychophysics*, **21**, 399-407.

Silverberg, R., Bentin, S., Gaziel, T., Obler, L.K., & Albert, M.L. 1979 Shift of visual field preference for English words in native Hebrew speakers. *Brain and Language*, **8**, 184-190.

Silverberg, R., Gordon, H.W., Pollack, S., & Bentin, S. 1980 Shift of visual preference for Hebrew words in native speakers learning to read. *Brain and Language*, **11**, 99-105.

Sloboda, J.A, 1976 Visual perception of musical notation : Registering pitch symbols in memory. *Quarterly Journal of Experimental Psychology*, **28**, 1-16.

Sloboda, J.A 1978 Perception of contour in music reading. *Perception*, **7**, 323-331.

Sloboda, J.A. 1982 Music performance. In D. Deutsch (Ed.), *The Psychology of Music*. Academic Press.

Sloboda, J.A. 1985 *The Musical Mind : The cognitive psychology of music*. Oxford University Press.

Spellacy, F. 1970 Lateral preferences in the identification of patterned stimuli. *Journal of the Acoustical Society of America*, **47**, 574–578.

Sperling, G. 1960 The information available in brief visual presentations. *Psychological Monographs*, **74** (Whole No. 498), 1–29.

Sperry, R. 1982 Some effects of disconnecting the cerebral hemispheres, *Science*, **217**, 1223–1226.

Springer, S.P. & Deutsch,G. 1985 *Left Brain, right Brain* (Revised Edition). W.H.Freeman & Company.

Sternberg, S. 1966 High–speed scanning in human memory. *Science*, **153**, 652–654.

Sternberg, S., Monsell, S., Knoll, R.L., & Wright, C.E. 1980 The latency and duration of rapid movement sequences : comparisons of speech and typewriting. In R.A.Cole (Ed.), *The perception and production of fluent speech*. Hillsdale,N.J. : Erlbaum.

Sternberg, S., Wright, C.E., Knoll, R.L., & Monsell, S. 1980 Motor programs in rapid speech : Additional evidence. In R.A.Cole (Ed.), *The perception and production of fluent speech*. Hillsdale,N.J. : Erlbaum. Pp. 507–534.

Stevens, S.S., Volkmann, J., & Newman, E.B. 1937 A scale for the measurement of the psychological magnitude of pitch. *Journal of the Acoustical Society of America*, **8**, 185–190.

Sundberg, J., Nord,L., & Carlson, R. 1991 *Music, Language, Speech and Brain*. Macmillan Press.

Thomas, I.B., Hill, P.B., Carroll, F.S., & Garcia, B. 1970 Temporal order in the Perception of Vowels. *The Journal of Acoustical Society of America*, **48**, 1010–1013.

Tulving, E. & Thomson, D.M. 1973 Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, **80**, 352–373.

Tulving, E. & Watkins, M.J. 1975 Structure of memory traces. *Psychological Review*, **82**, 261–275.

Umemoto, T. 1984 Verbal factors in memory. *Kyoto University Educational Science Report*, **30**, 149–159.

Umemoto, T. 1990 The psychological structure of music. *Music Perception*, **8**, 115–128.

Umemoto, T. 1993 Music cognition and scale schema. In C. Izawa (Ed.), *Cognitive Psychology Applied*. Hove and London : Pp. 217–223.

Umemoto, T. & Mikumo, M. 1988 Scanning tone sequences as a function of length of the sequences. *The 24th International Congress of Psychology.* Sydney.

Umemoto, T. & Mikumo, M. 1989 Retrieval a tone in a melody. Perceptual cognition and representative cognition in control of performance. *The Ministry of Education, Science and Culture, Grant-in-Aid for Scientific Research* (62450024), 8–37.

Umemoto, T., Mikumo, M., & Murase, A. 1989 Development of tonal sense, —— A new test on cognition of pitch deviation ——. *Human Developmental Research*, **5**, 155–174.

Umemoto, T., Takeda, M., Arimoto. M., & Kondo, Y. 1976 Memory for melodies and coding. *Proceedings of the 40th Congress of Japanese Psychological Association*, 589–590.

Umemoto, T. & Tokumaru, T. 1990 Categorical character of tone interval judgment. *Proceedings of the 102nd Convention of Kansai Psychological Association*, 14.

Vaughn, J. 1983 Saccadic reaction time in visual search. In K. Rayner(Ed.), *Eye movements in reading : Perceptual and language processes*. New York : Academic Press. Pp. 397–412.

Wagner, M.T. & Hannon, R. 1981 Hemispheric asymmetries in faculty and student musicians and non-musicians during melody recognition tasks. *Brain and Language*, **13**, 379–388.

Walker, R. 1978 Perception and music notation. *Psychology of music*, **6**, 21–46.

Walker, R. 1981 Teaching basic musical concepts and their staff notations through cross-modal matching symbols. *Psychology of music*, **9**, 31–38.

Warren, R.P. 1969 Auditory sequence : Confusion of patterns other than speech or music. *Science*, **164**, 586–587.

Watkins, A.J. & Dyson, M.C. 1985 On the perceptual organization of tone sequences and melodies. In P. Howell, & I. Cross, & R. West (Eds.), *Musical Structure and Cognition*. Academic Press. Pp. 71–119.

West, R., Howell, P., & Cross, I. 1991 Musical structure and knowledge representation. In P. Howell, R. West, & I. Cross (Eds.), *Representing musical structure*. Academic Press. Pp. 1–30.

Wickelgren, W.A. 1965 Acoustic similarity and retroactive interference in short–term memory. *Journal of Verbal Learning and Verbal Behavior*, **4**, 53–61.

Wickelgren, W.A. 1966 Consolidation and retroactive interference in short–term memory for pitch. *Journal of Experimental Psychology*, **72**, 250–259.

Wickelgren, W.A. 1967 Exponential decay and independence from irrelevant associations in short–term recognition memory for serial order. *Journal of Experi–mental Psychology*, **73**, 165–171.

Wickelgren, W.A. 1973 The long and the short of memory. *Psychological Bulletin*, **80**, 425–438.

Wickelgren, W.A. 1975 The long and the short of memory. In D. Deutsch & J.A. Deutsch (Eds.), *Short–term memory*. Academic Press. Pp. 41–63.

Wickens, D.D. 1972 Characteristics of word encoding. In A.W. Melton & E. Martin (Eds.), *Coding Processes in Human Memory*. New York : John Wiley & Sons. Pp. 191–217.

Wickens, D.D., Born, D.G., & Allen, C.K. 1963 Proactive inhibition and item similarity in short–term memory. *Journal of Verbal Learning and Verbal Behavior*, **2**, 440–445.

Wing, H.D. 1948 Tests of musical ability and appreciation. *British Journal of Psychological Monograph* (Supplement, **27**).

Woodworth, R.S., & Schlosberg, H. 1954 *Experimental Psychology* (Revised Edition). New York : Holt, Rinehart, & Wiston.

Yoshizaki, K. 1986 Effects of learning experiences on the shift of hemispheric functional asymmetry. *The Japanese Journal of Psychology*, **57**, 35–38.

Zimmer, H.D. & Engelkamp, J. 1985 An attempt to distinguish between kinematic and motor memory components. *Acta Psychologica*, **58**, 81–106.

Zimmer, H.D. & Engelkamp, J. 1989 Does motor encoding enhance relational information ? *Psychological Research*, **51**, 158–167.

****

Anderson, J.R. 1980 *Cognitive Psychology and its Implications*. Freeman and Company, San Francisco and London.

Butler, D. 1992 *The musician's Guide to Perception and Cognition*. Schirmer Books. A Division of Macmilln, Inc.

Critchley, M. & Henson, R.A. 1977 *Music and the Brain : Studies in the Neurology of Music*. William Heineman Medical Books Limited, London.

Eysenck, M.W. & Keane, M.T. 1990 *Cognitive Psychology : A student's Handbook.* Lawrence Erlbaum Associates, Hillsdale. Hove and London.

Goldmeier, E. 1982 *The Memory Trace : Its formation and its fate*. Lawrence Erlbaum Associates, Hillsdale, New Jersey.

Green, R.L. 1992 *Human Memory : Paradigms and Paradoxes*. Lawrence Erlbaum Associates, Hillsdale, New Jersey. Hove and London.

Gregg, V. 1986 *An Introduction to Human Memory*. Rougtledge & Kegan Paul Limited.

Hebb, D.O. 1972 *Textbook of Psychology* (3rd Edition). W.B. Saunders Company.

Hebb, D.O. 1980 *Essay on Mind*. Lawrence Erlbaum Associates.

Kalat, J.W. 1984 *Biological Psychology* (2nd Edition). Wadsworth Publishing Company.

Klatzky, R.L. 1980 *Human Memory : Structure and Processes* (2nd Edition). W.H.Freeman and Company, San Francisco and London.

Lachman, R., Lachman, J.L., & Butterfield, E.C. 1979 *Cognitive Psychology and Information Processing : An introduction*. Lawrence Erlbaum Associates, Hillsdale, New Jersey.

Loftus, G.R. & Loftus, E.F. 1976 *Human Memory : The Processing of Information*. Lawrence Erlbaum Associates, Hillsdale, New Jersey.

Matlin, M. 1983 *Cognition*. Holt, Rinehart and Winston. The Dryden Press. Saunders College Publishing.

Mayer, R.E. 1981 *The Promise of Cognitive Psychology*. W.H. Freeman and Company, San Francisco and Oxford.

Moore, B.C. 1989 *An Introduction to the Psychology of Hearing* (3rd Edition). Academic Press.

Norman, D.A. 1982 *Learning and Memory*. W.H. Freeman and Company, San Francisco and Oxford.

Parkin, A.J. 1987 *Memory and Amnesia : An Introduction*. Oxford : Basil Blackwell.

Pierce, J.R. 1983 *The science of musical sound*. Scientific American Books.

Radocy, R.E. & Boyle,J.D. 1979 *Psychological Foundations of Musical Behavior*. Charles C. Thomas Publisher.