REFERENCES

Multiple Components of the Perception of Musical Sequences: A Cognitive Neuroscience Analysis and Some Implications for Auditory Imagery

BERENICE CARROLL-PHELAN & PETER J. HAMPSON

University College Cork, Ireland

A neurologically plausible model of the auditory perception of musical sequences is proposed, and some implications are derived for auditory imagery. In line with a cognitive neuroscience approach, a componential analysis of the major functions required of auditory perception and imagery for musical sequences is first carried out, and a minimal model of auditory imagery and perception is outlined. The minimal model makes a clear distinction between the processing subsystems required for analyzing pitch and rhythm. It also incorporates an auditory buffer for the brief retention of relatively unprocessed auditory input, an attention subsystem that permits selective intake of relevant auditory information, a memory subsystem for melodic strings that combines the outputs of pitch and rhythmic analyses into a long-term, flexible representation of the musical sequence, and an associative memory system that links the results of perceptual analyses of music with extraperceptual (semantic and episodic) information about the musical string. The present theory builds on a recent model of music perception proposed by Peretz (1993) and also accounts for data from studies of patients with disorders of rhythmic processing, thus indicating the close relationship between auditory perception and imagery for rhythmic sequences, timing, and motor processes. Recent data are used to refine and extend the minimal model and provide an organizing framework for current laboratory work.

R EMARKABLE success has been achieved in the past 25 years in unravelling the structural and functional aspects of visual imagery (Cooper & Shepard, 1973; Farah, 1985; Finke, 1980; Hampson, Marks, & Richardson, 1990; Kosslyn, 1980; Morris & Hampson, 1983; Paivio, 1971), but only very recently has any substantive work emerged on auditory imagery. Examples include studies of songs (Halpern, 1988, 1989), sounds in the environment (Intons-Peterson, 1980; Segal & Fusella, 1970), musical instruments and timbre (Crowder, 1989a, 1989b; Pitt & Crowder, 1992), vocal

517

sound (Geiselman & Bjork, 1980; Geiselman & Glenny, 1977; Johnson, Foley, & Leach, 1988; Nairne & Pusen, 1984), detection of unfamiliar pure tones (Farah & Smith 1983), and unfamiliar musical phrases (Weber & Brown, 1986). Much of this recent work has, explicitly or implicitly, shadowed earlier work on visual imagery. For instance, Halpern (1988) has drawn parallels between an auditory task in which subjects scan imaged tunes for the presence of certain auditory segments and earlier visual scanning experiments (e.g., Beech, 1978; Kosslyn, Ball, & Reiser, 1978). Similarly, just as participants are thought to be able to transform visual images in various ways, with invariant aspects of images maintained under, say, rotation (Cooper & Shepard, 1973) or size changes (Kosslyn, 1975), so they are also able to transform auditory images, perhaps by mentally transposing a tune into a higher key (Halpern, 1992) or imagining it played by different instruments or at different speeds (Halpern, 1992; Pitt & Crowder, 1992).

Such demonstrations are interesting in their own right and undoubtedly add to our knowledge of the empirical properties of auditory imagery, but attempting to reproduce the early successes of visual imagery research is neither the only nor necessarily the most principled way to tackle the phenomena in question. An alternative approach is one that considers imagery as a phenomenon closely related to perception. In the visual domain, this has yielded two chief benefits. First, the componential structure of imagery has gradually been teased out. Thus, by careful use of interference methods, Baddeley and colleagues and others have shown visual imagery, like visual perception, to have both visual and spatial aspects (e.g., Baddeley & Lieberman, 1980; Logie & Baddeley, 1990). Second, the processing structures common to both imagery and perception can be identified, with correspondences established at various levels (Finke, 1989). Similar benefits might accrue from adopting the related working hypothesis for audition, namely, that auditory imagery is derived from auditory perception, and, therefore, depends in whole or in part for its reinstatement on the processing structures of the auditory perceptual system. Several recent authors are alert to the potential links between imagery and perception (e.g., Farah & Smith, 1983; Halpern, 1992; Pitt & Crowder, 1992; Weber & Brown, 1986), but they have stopped short of using the correspondence in a strong or principled way to derive an imagery theory out of a parent perceptual theory. It is this that we attempt in this paper, by specifying the major characteristics of such a perceptual theory, indicating, where appropriate, how it could then be used to generate a theory of imagery.

Further constraints on explanations of auditory imagery can be imposed by adopting a cognitive neuroscience perspective. Cognitive neuroscience brings to bear a broader set of data, from the brain sciences and from

neuropsychology, on theories previously constrained only by data from laboratory experiments on unimpaired participants. It also seeks accounts grounded in known aspects of neurologic structure and function, and its theories can, in turn, be computationally modeled (Kosslyn & Koenig, 1992). Once again, it is interesting to review progress on visual imagery as a result of this approach. Such a review indicates a sharpened interest in the use of neuropsychological data (e.g., Farah, 1984; Riddoch, 1990), brain imaging techniques (e.g., Goldenberg, Podreka, & Steiner, 1990; Marks, 1990), and the construction of models believed to correspond to known neurologic systems (e.g., Kosslyn & Koenig, 1992; Peretz, 1993).

In this article, we explicitly apply a cognitive neuroscience perspective to auditory perception and, following on from this, derive certain hypotheses on the nature of auditory imagery. We show, by accepting the working hypothesis that auditory imagery relies on the auditory perceptual system, and drawing largely on existing data, that it is possible to construct a neurologically plausible model that captures key aspects of the perception and imagery of musical sequences. In accepting the correspondence between imagery and perception, we do so in a broad rather than a narrow way and leave open the possibility that both perception and imagery are affected extensively by cognition (Intons-Peterson, Russell, & Dressel, 1992; Neisser, 1972). At this stage in the model's genesis, we limit our account to musical stimuli. We defend this restriction on the grounds that such stimuli can be well-defined and unambiguously reproduced. Unlike words, they can be selected to have minimal semantic content; they are, nevertheless, reasonably complex and still have ecological validity, and people often claim to have auditory images of musical stimuli. Despite this current restriction, we anticipate that a more generally applicable model of auditory phenomena could be constructed along the lines we propose.

We also show how such a perspective reveals aspects of auditory imagery thus far uncovered by existing research programs, namely its incorporation of rhythmic as well as pitch components, and a long-term memory for melodic strings. We do this by building on a recent model of music perception that already captures these distinctions (Peretz, 1993) and by incorporating data from studies of patients with disorders of rhythm processing.

In line with the cognitive neuroscience approach, a componential analysis of the major functions required of auditory perception and imagery for musical sequences is done first. Next, we outline a minimal model of auditory imagery and perception, explaining how such a model can account for the perception, recognition, and recall of musical sequences, and provide an organizing framework for current laboratory work. Relevant data on major aspects of the model, especially the pitch and rhythm subsystems,

are then examined. These data are used to refine and extend the minimal model and to show how a theory of imagery is derived from the parent perceptual theory.

Componential Analysis and a Minimal Model

Consider someone sitting in a waiting room who hears a familiar tune, say a passage from Beethoven's *Moonlight* Sonata, which they recognize, identify, and later in the day find themselves humming to themselves. As with other perceptual tasks, a number of logically distinct and distinguishable activities are involved here. Listeners must first detect and attend to the stimulus and retain sufficiently large chunks to apprehend its larger structure. Second, they must be capable of performing some preliminary analysis on the input sufficient to allow them to recognize the piece as familiar. Third, with a familiar tune that they not only accurately perceive but go on to identify, they must also be capable of accessing further semantic and verbal information about the piece, perhaps remembering its composer, its musical form, its name, and so on. They may, in addition, be able to summon up specific autobiographical memories of hearing the piece in a concert hall or while relaxing at home. Finally, they may be able to reproduce aspects of the initial perceptual experience at a later date.

Logically speaking, the processing operations after the initial registration of the stimulus needed to accomplish all this, need not necessarily occur in a fixed order, need not all be mandatory, and (given the findings in other areas of perception) are likely in practice to be highly interactive. For instance, a person might retrieve the name of the composer of a piece of music, before its title, or might recognize an unknown tune as exhibiting sonata form. Nevertheless, some ordering of stages would seem to be likely, and so we assume that initial stimulus registration is typically followed by processing of manageable chunks of information in which properties of the piece are extracted, sufficient to cue a representation of the melody in question in a tune-specific memory system. Activation of a stored representation will in turn lead to coactivation of information in a general purpose associative memory, allowing semantic, episodic, autobiographical, and emotional information to be related to the stimulus (cf Kosslyn & Koenig, 1992). Such information may also assist in the identification and interpretation of variously sized chunks of a musical sequence (melody).

It is interesting to compare our present sketch of the major psychological functions needed for music perception with the four dimensions of musical stimuli outlined by Umemoto (1990). Umemoto distinguishes between the dimensions of (1) music as sound, (2) music as an object of perception, (3) the compositional structure of music, and (4) the composi-

tional content of music. The scheme ranges from the microanalysis of single tones to the macro empathic response to the entire music piece. Psychological processes involving sensation, perception, cognition, and evaluation are assumed to deal with these dimensions (see Williams, 1982, for a similar analysis). Thus, "music as sound" refers to characteristics of a single tone, such as its pitch, loudness, timbre, or duration. We assume this basic level of representation to be available in the early stages of stimulus registration. "Music as an object of perception" refers to the perception of melody, harmony, and rhythm over sequences of tones. It is this dimension that we assume to be supported by the perceptual processing and tune memory subsystems. "Compositional structure" involves comparisons of similarity between entire musical phrases or parts of the composition. We see this as a higher order skill involving recursive, relational processing between current perceptual activity and earlier results integrated into the tune memory subsystem, possibly supported by additional knowledge brought to bear from associative memory. Finally, the "compositional content" of a piece involves "the emotional cognition or empathic understanding of the meaning, character or script of a composition." (Umemoto, 1990, p. 117). This, we assume, involves the triggering of the powerful full resources of associative memory by the tune memory system. A schematic outline of the functional components identified thus far and their correspondence with the dimensions identified by Umemoto are depicted in Figure 1.

In our basic model, we assume that perceptual processing is typically data-driven, but this neither rules out the possibility of concept-driven support at higher levels nor excludes cognitive factors in segmenting the flow of perceptual information. We acknowledge the point made by Williams (1982) that the ongoing dynamic character of music locks the listener into a psychological loop, strongly controlled by their past experience with music. We agree that "The listener's ability to decode raw neural signals, into sensory percepts and then subsequently organize those percepts is dependent on the extent to which their musical or extra musical schemas allow them to decode, classify and order sonic events" (Williams, 1982, p. 35).

FUNCTIONAL COMPONENTS		MUSIC DIMENSIONS
input		
registration and attention	}	music as sound
perceptual processing	}	music as perception
tune memory subsystem	}	compositional structure
associative memory	}	compositional content

Fig. 1. Hypothesized major functional components of auditory perception and their relation to four musical dimensions.

Williams goes on to state that "There is a critical ratio between the degree to which past experience permits the rate of processing to match the rate of music information; insufficient schemas result in information being discarded or lost." (Williams, 1982, p. 35). In terms of our framework, the presence of rich organizing structures in associative memory will allow more and more detailed comparisons to be made within the tune memory subsystem. In turn, greater expertise in dealing with the dimension of music content will allow more sophisticated comparisons to be made of compositional structure (cf Gromko, 1993, for a related analysis). However, although the richer associative memories in the minds of experts are likely to facilitate their music processing, we also expect that musical expertise will turn out to be multifaceted, like many cognitive abilities, and to depend on a individual variation in a wide variety of perceptual and postperceptual stages.

Issues of expertise aside, this basic perceptual framework can be readily used to account for auditory imagery. Imaging a tune at some future date implies the ability to retain aspects of the previous interaction with the musical sequence and to recreate psychological conditions that resemble those produced by perception of the original sequence. One way in which this could take place is via the replay or reprojection of information through the system. For the moment, we do not wish to speculate on the precise details of this process, but wish simply to draw a possible analogy between the reactivation of tune-specific information that can be "replayed" back through the auditory system and subjected to an analysis similar to that carried out in perception, and the way in which visual imagery is thought to rely on visual perceptual processes (Finke, 1985). Thus, as we indicated earlier, we adopt as a strong working hypothesis the notion that auditory imagery makes use of major subsystems involved in the intake and coding of auditory perceptual input, partly on the basis of existing evidence showing modality specific interference between auditory imagery and perception (e.g., Intons-Peterson, 1980; Segal & Fusella, 1970; Serafine, Crowder, & Repp, 1984), also because of neuropsychological data indicating the correspondence between auditory imagery and perception (Zatorre & Halpern, 1993), and partly on the pragmatic grounds that this guiding hypothesis has proved its utility in the visual domain.

Further clues to the nature of perceptual systems in general and music perception in particular can be gleaned by a closer examination of the type of stimuli with which they have to deal. With musical stimuli, an important psychological distinction can be made between properties of the stimuli themselves on the one hand, and interstimulus qualities on the other. This distinction is of course well captured by the first two musical dimensions outlined by Umemoto (1990), but it deserves closer consideration. Audible musical stimuli, like all auditory phenomena, occur in nature as sound waves. Physical variation in sound energy carries important psychological infor-

mation on the pitch, timbre, loudness, and duration of the notes in question. However, a description of the qualities of individual musical stimuli, although necessary, is clearly insufficient for a complete theory of music perception. Just as a theory of word perception alone, without a theory of grammar, is inadequate as an account of sentence processing, inadequate too is a theory of music perception that deals only with individual notes. Important interstimulus qualities must be considered. These include tonal relationships, and, crucially, the temporal organizing qualities of beat, rhythm, and tempo. Of these stimulus and interstimulus qualities, we treat pitch and rhythm relationships as primary because these two dimensions uniquely combine to produce a melodic string. Such a melodic string can then be thought of as exhibiting musical constancy despite changes in timbre, loudness, tempo, and so on. Note that we are not requiring such a string to be aesthetically pleasing nor acceptable as music. Borrowing terms from the grammarians, we accept well-formed and ill-formed strings as potential material for auditory perception and here draw attention merely to their uniqueness as stimuli and to their compositional character.

Assuming that rhythm and pitch singly and jointly specify important aspects of the incoming information, it follows that these dimensions must at some point be extracted after immediate stimulus registration. Just as visual and spatial qualities are extracted during visual perception, we can hypothesize that after the initial registration of a musical stimulus, pitch and rhythmic dimensions are analyzed at some point in the processing stages after musical input. However, we also submit that, just as in vision, where the system must keep track of the correlation of visual and spatial dimensions to allow the perception of objects-at-locations, so too the auditory processing system must, at some stage, link up its pitch and rhythmic analyses to permit the perception of temporally organized pitch sequences or melodic strings. We further assume that pitch and rhythmic analyses are done simultaneously in parallel on the incoming string. On a priori grounds alone, completely serial processing of pitch and rhythm would most likely be too slow to keep up with the processing of even moderately paced musical strings, let alone the rapid changes of, say, an arpeggio. Other processing arrangements are possible as alternatives to strict serialism. For instance, the auditory system could process sections of extended pitch sequences in parallel and then pass these down line to the rhythmic subsystem for further analysis. Although this arrangement would undoubtedly cut down the number of processing steps, maintaining the temporal correspondence between pitch and rhythmic analyses would then be more complex than with fully parallel processing, so we assume for now that fully parallel processing of pitch and rhythmic sequences occurs.

The utility of the assumption that rhythm and pitch are processed in parallel is borne out by recent empirical work of Peretz and Kolinsky (1993), who have shown that pitch and rhythm dissociate in patients with brain

injuries. Peretz(1993) has outlined a model of music processing that is similar in core components to our own and is designed to account for data from brain-damaged patients and healthy people. Like ourselves, Peretz also proposes parallel structural analysis of pitch variations and temporal (rhythmic) features. Other laboratory studies have obtained results compatible with the parallel route model of pitch and rhythmic processing (Halpern, 1984; Monahan & Carterette, 1985; Monahan, Kendall, and Carterette, 1987; Palmer & Krumhansl, 1987a, 1987b). Furthermore, in the neuropsychological literature, several other cases have been described in which melody and rhythm were separately affected by brain damage. For example, Mann (1898, cited in Dorgeuille, 1966) showed that rhythm may be spared when melody is lost in terms of singing ability, and vice versa (Brust, 1980; Mavlov, 1980). In reading music, a similar double dissociation was found for rhythm by Brust (1980) and for melody by Dorgeuille (1966) and Assal (1973).

There is, however, a further body of evidence in the current literature supporting the alternative view that melody and rhythm are not psychologically independent and are fully integrated in both music perception and memory by listeners to musical sequences (Boltz, 1989a, 1989b, 1989c; Boltz & Jones, 1986; Jones, 1987; Jones, Boltz, & Kidd, 1982). Reviewing this evidence, Jones and her colleagues use the term "dynamic pattern structure" to refer to the ways in which rhythm and melody combine. Jones uses the term "dynamic" to emphasize the integration of temporal relationships with other aspects of musical pattern structure, in particular melody and rhythm. She argues that pitch relationships in a musical pattern contribute to its melody and harmony, and time relationships to meter, rhythm and tempo, but that both qualities are intimately related. Her analysis of melodic and rhythmic interplay focuses on what she calls a "joint accent structure" that results when both melodic and rhythmic accent patterns uniquely combine and mutually cohere (Jones, 1987, p.625).

Jones and her coworkers provide a strong argument, with supporting experiments, against the separability of rhythm and melody, which it is not our intention to ignore. However, the fact that such integration can take place in normal listeners to musical sequences does not imply that there cannot be earlier separation of the processing of pitch sequences (rather than melodies) and rhythm. To take an analogy from vision, local analyses of object and location information can be carried out independently yet still result in an integrated percept of a scene in which the two are combined. Similarly, in audition, the separate and parallel processing of pitch and rhythmic sequences can be followed in the perceptual sequence by their reintegration into a rhythmically organized melody.

Some further aspects of the general perceptual framework can also be clarified at this stage. From observation and empirical data, it is clear that

individuals are capable of retaining memories for tunes and melodic sequences for long periods of time (Bartlett & Snelus, 1980). The ability to reproduce a tune from memory indicates that both its major dimensions, namely its rhythm and pitch sequences, must be held in some durable form. Whether pitch and rhythm are stored in an integrated form or are represented separately and recombined at retrieval is an issue we will examine later. What we argue for now is that the underlying tune memory is likely to retain the invariant aspects of a melodic sequence, irrespective of particular instantiations of the sequence in a given key or whether played by a particular instrument or sung by a particular voice. On logical grounds alone, such a comparatively stimulus-independent representation appears necessary to accomplish the common everyday feat of recognizing the same tune even though it may be played in an unfamiliar way, say by a piano instead of a guitar, or in a higher key than usual (but see Crowder, 1989a, 1989b). A further aspect is that an auditory buffer, or fairly transient store, will be required to retain musical sequences in a direct, quasi-sensory form to allow sufficiently large temporal segments to be selected by processes of auditory attention and relayed for further analysis. This follows from the assumption that both pitch and rhythm processing are likely to operate over groups of stimuli, with the former dealing with contour, interval, and tonal pitch relationships and the latter dealing with durational properties of musical patterns. Finally, auditory attentional processes are assumed to modulate or control the flow of information from the buffer.

   Adding these considerations into the basic framework leads to the extended model depicted in Figure 2.
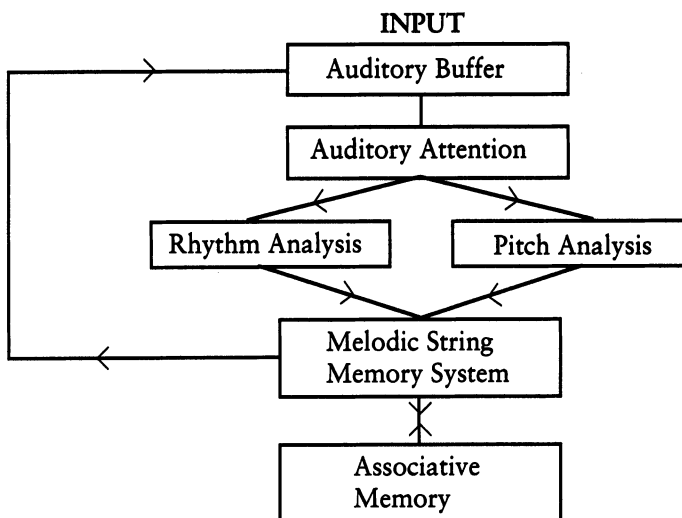

Fig. 2. An extended model of auditory perception and imagery.

Empirical Support and Elaboration of the Model

So far we have developed, in outline, a basic model of auditory percep-
tion for musical sequences, largely, on logical grounds following a basic
analysis of the high-level computational requirements of the task. In this
section, evidence in support of the proposed model is considered. The evi-
dence reported falls into three broad categories: data from cognitive psy-
chological experiments, data from investigations of the impaired and in-
tact abilities of patients with brain injuries, and data from neurophysiological
and neuroanatomic studies of brain structure and function.

Aspects of the model that are reasonably well established in existing
accounts of auditory perception are the auditory buffer, auditory attention,
and associative memory (cf Kosslyn & Koenig, 1992). Given that these
components are relatively uncontroversial, we deal only briefly with them
here, indicating where we see contact with other theories and existing data.
Aspects dealt with in more detail are the pitch and rhythmic analysis sub-
systems and the melodic string memory system. As we mentioned earlier,
these have been treated separately, often by independent groups of research-
ers, with little attention until very recently paid to their interrelationships.
To our knowledge, only Peretz and her coworkers have examined both the
dissociability of pitch and rhythm and their potential links with a melodic
memory system, or "tune lexicon" as they have named it, in any detail
(e.g., Peretz, 1993; Peretz & Kolinsky, 1993). We concur with Peretz that
there is more known about the pitch route in the processing of musical
information than the route via rhythm, but we take this opportunity to
show that there is a body of knowledge on impairments of rhythm that can
be marshaled in this context and that implicates brain areas in music per-
ception and memory other than those generally considered to be involved
in auditory perception and imagery.

AUDITORY BUFFER

Good evidence exists for a system that can retain auditory sensory infor-
mation for brief periods of time in a relatively unprocessed form. Neisser
(1967) coined the term *echoic memory* to refer to such a system, arguing
that it was the auditory equivalent of visual iconic memory. We conceive of
this as the first stage of the auditory memory/processing system, with a
duration of only several hundred milliseconds, and for information to be
held by it in an acoustic form (e.g., Engle, Cantor, & Turner, 1989; Penney,
1989). We further assume that processing at this early stage is domain
general, that is, the auditory buffer/echoic system can support both musi-
cal and nonmusical auditory stimuli, including speech. We distinguish be-
tween such a sensory store and later components of the system that accept

information from it and retain information for longer periods. Thus, we are not suggesting that the auditory buffer is the equivalent of, say, the phonological store in a working memory system, but we do suggest that working memory subsystems and other auditory processing subsystems such as those for pitch and rhythm discussed later will take their inputs from the auditory buffer (see Baddeley, 1990, for a similar view on the relationship between working memory and echoic memory).

It appears highly likely that the neurophysiological locus of the echoic system is to be found in the primary auditory reception areas (Kosslyn & Koenig, 1992). These are known to be tonotopically organized, as electrophysiological and brain scan studies have shown, with stimulation resulting in the individual hearing tones of a particular pitch, maybe also perceived as coming from a particular direction (Penfield, 1967). Furthermore, lesions in these areas can result in loss of the ability to discriminate particular sounds (Heffner & Heffner, 1990, 1986; Mendez & Geehan, 1988). Also, neural activity in these areas can be related to subjectively experienced psychological dimensions such as duration. For example, recent magnetoencephalographic work by Lu, Williamson, and Kaufman (1992) shows that the duration of neuronal activity in the primary auditory cortex can be used to predict the psychophysically determined duration of memory for loudness of a tone.

Echoic memory, as captured by the auditory buffer, permits a basic, sensory representation of the stimulus to be retained for sufficient time for further analyses to take place. In our model, no other features of echoic memory are postulated at this stage.

ATTENTION SUBSYSTEM

This is not the place to review fully the classic literature on auditory selective attention or to become embroiled in the debate concerning the absolute locus of any selective activity. Suffice it to say that it has been amply demonstrated that participants are able to focus selectively on one auditory input in preference to other competing inputs and that inputs clearly distinguishable on the basis of physical cues afford more effective selection than inputs that are physically similar (e.g., Broadbent, 1958; Cherry, 1953). In the case of linguistic inputs and in the absence of clear physical cues, selection using higher order cues is possible (e.g., Treisman, 1964), although the "later" the selection, the more effortful it appears to be (Eysenck & Eysenck, 1979; Johnston & Heinz, 1978).

As is well known, selective filters or hierarchies of tests through which incoming information must pass were frequently posited to explain the results of classic dichotic listening experiments (Broadbent, 1958; Treisman, 1964). However, despite their attraction, the neurophysiological plausibil-

ity of such devices remains doubtful. Consequently, it is useful to remember that non–filter-based accounts of attention have also been proposed (Neisser, 1967, 1976). Using a visual analogue of shadowing, Neisser and Becklen (1975) showed that participants were able to attend selectively and successfully to one of two spatially overlapping videotaped sequences, while noting little of the unselected sequence. Emphasizing the active nature of the perceptual following of an action sequence, rather than passive filtering, they concluded that attention is a by-product of skilled perceiving, not an additional blocking process, and claimed that unselected sequences simply went unprocessed. "Selection as skilled perceiving" is also a useful way to explain the results of a recent experiment on auditory interleaving by Dowling (1973a), an analogous experiment, in the musical domain, to Neisser and Becklen's study. Dowling used two familiar tunes, "*Frère Jacques*" and "Twinkle Twinkle," temporally interleaved, to demonstrate salient features of music attention. The two melodies when played together, spatially separated, were easily distinguishable from each other played at the rate of 7.5 notes per second, using pitch cues as the basis for selection. However, as with Neisser and Becklen's visual study, participants were able to attend to only one melody at a time. In a further example, Dowling allowed the pitch ranges to overlap but changed the timbre of each melody. Once again, and in line with classic work on attention, the physical cue of timbre was an effective cue for selection between stimuli, as were loudness differences and spatial separation of the tunes.

In the present model, we assume that interactions will occur between patterns activated in the melody memory system and those in the auditory buffer, resulting in selective attentional emphasis of those inputs whose structure can be predicted from knowledge of the sequence in question. Such highlighting will result in the orientation of processing subsystems toward the relevant musical stimuli and successful match of the input with long-term patterns in an interactive-activation fashion (cf McClelland & Rumelhart, 1986).

PITCH ANALYSIS SUBSYSTEM

In the current model, the functions of the pitch analysis subsystem are (1) to accept relatively unprocessed information from the auditory buffer, (2) to analyze this information with particular reference to the interrelationships that exist within and between sequences of notes, (3) to communicate the results of such analysis to other parts of the system, especially those subsystems responsible for maintaining a long-term memory representation of an entire melodic sequence.

To spell out more clearly the properties of this putative subsystem, we first consider the various physical and associated psychological distinctions

that can be made about pitch. Next, and in the light of these distinctions, we review the performance characteristics of human participants in studies of pitch processing together with relevant neuropsychological evidence. We then indicate what we consider the pitch analysis subsystem to be capable and incapable of doing, and in so doing delineate it from other subsystems, notably the auditory buffer on the one hand and the melodic memory system on the other.

Three major aspects of musical pitch are interval, scale, and contour. Interval refers to the distance in pitch between notes, which in turn depends on their frequency in hertz. In Western music, the smallest interval is a semitone, which represents a frequency difference of around 5.9%. A scale is a group of notes subject to certain pitch constraints. The first constraint is that notes in a scale must be discriminable in pitch from one another when played in succession. Second, scales must exhibit octave equivalence, that is, notes an octave apart, for example two Cs are perceived by most listeners as similar. Western musical tuning further divides an octave into 12 semitones known as the chromatic scale. Using this, any melody can be transposed so as to start on any note and be reproduced without distortion by preserving the appropriate intervals between notes. Thus, musical transposition alters the overall place in which the melody is hung, with the music sounding in general higher or lower in pitch, but preserves the relative interpitch intervals. In practice, although division of each scale into 12 semitones is possible, typically only around 7 or 8 notes per scale (the tonic sol-fa) are used in music. Finally, sets of notes can be characterized in terms of *contour*, the pattern of ups and downs of pitch in a melody, together with additional information on where they are to be "hung" on a scale (Dowling, 1973b).

Each of these pitch dimensions has psychological validity and performance implications. Interval and contour, in particular, can be shown to have distinguishable effects on matching pitch sequences (Deutsch, 1980; Deutsch & Feroe, 1981; Francès, 1958; Jones, 1974) and to relate to the level of skill of the participants. Thus, Bever and Chiarello (1974) argued that whereas skilled musicians rely more on intervals when recognizing melodies, nonmusicians rely more on contour. Furthermore, they suggested that whereas interval judgments are carried out by left-hemisphere processes, contour judgments depend on right-hemisphere processes. This distinction and its hemispheric locus receives further support from work by other researchers, chiefly Peretz and colleagues on nonmusicians (Peretz, 1987; Peretz & Morais, 1987; Peretz, Morais, & Bertelson, 1987) and on musicians (Peretz, 1993; Peretz & Babai, 1992). Converging evidence also comes from studies of persons with brain injuries by Peretz (1990), who has teased apart the laterality effect further. She has shown that left-hemisphere lesions do not affect patients' abilities to discriminate and recognize

melodies in terms of their contour but do disrupt their ability to discrimi-
nate and recognize melodies using intervals. Lesions in the right hemisphere,
conversely, disrupted both interval- and contour-based procedures. Thus,
whereas a *dissociation* after left-hemisphere damage can be demonstrated
between impaired interval-based performance and relatively intact contour-
based performance, an *association* is demonstrated after right-hemisphere
damage that results in impairment to both. This pattern of results leads
Peretz to conclude that contour extraction, by the right hemisphere, is a
necessary step before the more detailed processing of intervals by the left
hemisphere (Peretz, 1993). It remains to be seen whether this detailed divi-
sion of labor between the hemispheres is supported by further work, be-
cause to date most studies of the neurologic localization of pitch have gen-
erally implicated right-hemisphere processes in the temporal cortex (e.g.,
Milner, 1962; Samson & Zatorre, 1988; Zatorre, 1985; Zatorre & Halpern,
1993). However, a recent positron emission tomography (PET) study by
Zatorre, Halpern, Perry, Meyer, and Evans (in press) has shown bilateral
activation to occur in the auditory cortices during the perception and im-
agery of familiar songs. Although not incompatible with the idea that pitch
processing relies on activity in both cerebral hemispheres, it should be noted
that their task is more complex than ones typically used to study pitch
processing alone (e.g., Deutsch, 1969, 1970, 1972, 1973) and includes rhyth-
mic and verbal components as well. (Their participants had to judge pitch
differences between two words in imaged or perceived familiar songs). The
fact that other brain areas were active during the task, including the frontal
lobes, the left parietal lobes, and supplementary motor areas, further sug-
gests that their task, as they acknowledge, involves more than simple pitch
processing. Activation in some of these areas is thus, as we shall see shortly,
easier to explain within the larger model of musical processing, rather than
the pitch analysis subsystem alone.

Evidence also exists that individuals are sensitive to scale and its associ-
ated tonal functions. By "tonal functions," we mean the rules governing
the set of pitches and their combinations generally acceptable in Western
tonal music, and we use "tonal knowledge" to refer to the psychological
structures, schemata, processes, or representations needed to interpret tonal
functions. The notion of tonal functions and associated tonal knowledge
suggests that there are clear, fixed, construction rules governing musical
tonal sequences and that, like a grammar in natural language, such "rules"
should specify whether musical sequences are deemed to be well formed or
poorly formed. However, it may prove difficult in practice to specify pre-
cisely what constitutes a well-formed sequence, even within the Western
musical system, given its wide variation in musical forms and genres. For
example, an unacceptable transition within a piece by Mozart may be
deemed perfectly permissible if it appears within a piece by, say, Stockhausen.

Cultural, historical, and artistic factors will strongly influence what is deemed to be musically well formed. More generally, the Western attunement to the tonic sol fa reflects the unique development of music in one cultural nexus. Music from the Middle Ages, or, say, traditional Japanese or Indian music, acculturates different judgmental skills in individuals frequently exposed to such musical forms (Castellano, Bharucha, & Krumhansl, 1984). We are also conscious that these abilities are likely to vary as a function of an individual's perceptual skill and musical sophistication. Unlike language, where a grammaticality judgment can be made by appeal to the basic intuitions of a native speaker, more careful specification of the characteristics of musicality judges and judgments appears to be necessary. Therefore, it may be safer to restrict the notion of well-formedness and consequent tonality as affecting judgments made within a related category of music, by an individual with a particular level of skill and experience. Thus, although it may be difficult to make absolute or universal statements about tonal functions and knowledge, these constructs still have psychological validity.

Evidence to date suggests that tonal knowledge can be demonstrated by: the relative ease with which individuals detect changes in tonal as opposed to atonal melodies (Francès, 1958); the inability of subjects to make interval-based judgments in atonal melodies (Dowling & Fujitani, 1971); their contrasting ability to distinguish interval changes in a tonal melody when followed by atonal, contour-preserving sequences (Dowling, 1978); and their skill in predicting terminal from earlier portions of melodies (Bharucha, 1987). Sometimes this tonal knowledge can be very precise. For example, Halpern (1989) has demonstrated that people can generate or select the starting note of a tune after they generate the tune inside their heads, and she has shown that people have quite good representations of the absolute pitch of familiar songs.

Neuropsychology provides additional support that tonal knowledge is not only central to pitch processing, but is also dissociable from contour and interval processing. Following examination of her patient, GL, Peretz (1993) concludes that the patient was able to make use of contour and interval information but was insensitive to tonal structure.

We have looked at the three major psychological distinctions, interval, contour, and tonal functions, that may be made about musical pitch. A further aspect that we now briefly note is timbre, which refers to the "color" or characteristic sound of a given note played by a particular instrument. Thus, an A played on the piano, violin, or trombone may share the same pitch but differ in timbre. Also, whereas pitch characteristics such as contour, interval, and tonal structure depend on sets of notes, timbre can be a characteristic of single notes or tones. It is not our intention here to review in detail work on timbre, but simply to acknowledge that this dimension too can be psychologically distinguished from others such as interval, scale,

and contour (Crowder, 1989a, 1989b), imaged (Pitt & Crowder, 1992) and potentially impaired through brain injury, especially in the right temporal lobe (Samson & Zatorre, 1994).

Thus far, we have reported evidence for interval, contour, tonal, and timbre processing abilities and their possible impairment. Experiments on these abilities have typically been conducted with simple pitches or pitch sequences, but, of course, music includes more complex harmonic relationships of consonance and dissonance between pitches. Higher order perceptual judgments can be made along these dimensions too, and these may depend on systems that are neuropsychologically distinguishable from those needed to process more basic aspects of pitch such as tonal functions. Distinguishing between different sorts of pitch tasks in this way, Tramo, Bharucha, and Musiek (1990) examined the music perception and cognitive ability in a 30-year-old man, after damage to the auditory cortex caused by strokes. Complete bilateral lesions of the transverse gyri of Heschl, together with partial bilateral lesions of right and left superior temporal gyri, in their patient, seriously impaired consonance perception, resulting in major triads judged to sound dissonant. Conversely, their patient had normal pure-tone sensation thresholds, and associative priming of spectral judgments by harmonic context was also unimpaired. They assumed the latter to be a cognitive task involving the hierarchical structuring of pitch information. Tramo et al. suggest that sensory and cognitive functions involving tonal information processing of music, left intact by damage to the cortical areas in question, are dissociable from functions supporting consonance perception.

From this brief review of the literature on pitch processing and its impairments, we can conclude that the pitch processing system is sensitive to interval, contour, scale, and tonal functions and timbre, and it can be used to make higher order judgments such as those involving consonance, well formedness of sequences, and perceptual expectations.

Within this system, little is yet known in detail about how pitch is represented, although current models (e.g., Shepard, 1982a, 1982b) do go some way toward explaining certain aspects. In Shepard's model, pitch is assumed to be represented in a helical fashion, an arrangement that not only captures the similarity of notes near in pitch to one another, but also octave equivalences and other tonal relations to which individuals are sensitive (cf Krumhansl & Shepard, 1979). Nor is much known yet about the dynamics of the system that is needed to deal with sets of incoming notes and to perform high-level harmonic analyses, although some progress has been made in understanding how well-formedness judgments are characterized. Notice, first, that such judgments can presumably be carried out even though individuals do not know (in the sense of having previously experienced) the precise string in question. This suggests that long-term representation

of the precise string need not be involved, but that some ability to make a generative judgment is. Despite our earlier caveats, it is instructive to draw an analogy here with finite-state grammars (see, Miller 1967, for discussion). Musical sequences can be thought of as finite-state strings in which later portions of a sequence are partially determined by earlier portions. This further suggests that a skilled listener, capable of detecting violations of sequence, is able to use the predictability inherent in a string. Nor need the likening of musical strings to grammars commit us to a model of pitch judgments in which stored rules or grammars are deployed (but see Deutsch & Feroe, 1981; Lerdahl & Jackendoff, 1983, for examples of the latter). Hubbard and Stoeckig (1992) review various alternatives including connectionist models that capture key intertonal and harmonic relationships but, as currently constituted, these models deal only implicitly with melodic sequences.

Finally, we distinguish the pitch analysis subsystem from the auditory buffer on the basis that the former engages in active perceptual analysis of pitch strings by contrast with the passive registration of sensations by the buffer. We also delineate the pitch analysis subsystem from the melodic memory system, which we consider shortly. We do so largely on conceptual grounds, accepting that current psychological and neuropsychological evidence in support of the division is still unclear. According to our view, the pitch analysis system is distinguishable from the melodic memory system on the basis of three criteria. First the pitch analysis system is used for the processing of any musical sequence, known or unknown; it is not used to recognize sequences that have been previously heard or to categorize melodies as novel. Like Peretz (1993), therefore, we distinguish between the process of *access* to an entry in a musical memory system or lexicon and the *activation* of the entry in the lexicon, the former, in our model, being that carried out mainly by the pitch analysis system, the latter by the melodic memory system. Our reasons for adopting this view are largely based on a meta-hypothesis that music processing will be found to behave in a way similar to other perceptual activities such as reading (e.g., Morton, 1969) or face recognition (e.g., Bruce & Young, 1986), where distinctions between access and memory entries have already been fruitfully applied. This distinction does not commit us to any particular model of musical access. Although for the moment we still treat this as a data-driven process, we suspect that interactive activity may well occur between the melodic memory system and the process of access itself. (We also assume, and concur with Peretz, that access via the pitch system is more reliable than access through the rhythmic subsystem; cf Peretz, 1993). Second, we assume that analyses in the pitch processing system are carried out in a weakly modular way, independently from rhythmic analyses, whereas (as we shall see) a full entry in our melodic memory system includes both rhythmic and pitch com-

ponents. Third, we associate the activities of the pitch processing system more with the short-term or working memory activity involved in the on-line processing of comparatively short musical sequences, whereas the melodic memory system is involved in the acquisition, retention, and reactivation of entire melodic sequences and general tonal knowledge. This "memory criterion," which imposes key constraints on the activities of the pitch subsystem, results from the temporal and potentially rapidly changing aspects of the musical stimulus and the short storage time of the auditory buffer that feeds information, through the auditory attention subsystem, to the pitch processing system. Thus we hypothesize that only local not global analyses can be carried out by the pitch subsystem. In line with this suggestion is the discovery by Zatorre and Samson (1991) that unilateral temporal lobe lesions can impair the retention of pitch information in short-term or working memory, suggesting the existence of neural systems specialized for this purpose. Further evidence for a specific, tonal working memory system, and indications that, depending on the task, it may rely to a greater or lesser extent on frontal lobe structures comes from recent PET studies of tone-monitoring tasks (Petrides, Alivisatos, Meyer, & Evans, 1993) and pitch judgments with novel melodies (Zatorre, Evans, & Meyer, 1994). In these studies, increased cerebral blood flow was recorded in area 46, with rightward asymmetry in both cases. Comparisons of earlier portions of a long familiar melody with a current portion, by contrast, would seem logically to require a more durable form of representation to be used in conjunction with current information than the forms used in dealing with novel melodies. In cases such as these, the focus of cerebral blood flow shifts posteriorly to the boundary of areas 45 and 9 and possibly area 44 (Zatorre et al., in press). However, as we stated earlier, these tasks go beyond simple, on-line pitch processing and also require access to long-term memory representations in what we have termed the melodic memory system. To take an analogy from reading, pitch processing, as we see it, is more akin to the ongoing perception of single words and short word groups, rather than the longer term recognition, identification, comprehension, and ultimately memory of larger textual units such as paragraphs and stories.

We defend our separate treatment of pitch and melodic memory systems on the grounds we have just outlined, but also note that there are testable implications of the divide. If a separate tune lexicon or melodic system exists, various neuropsychological conditions could in principle arise that would not be predicted from a single-system view. For example, and most obviously, it would be possible in principle to discover disorders producing damage to the melodic memory system, which as a consequence leave pitch processing fully intact but affect the identification of melodies as old or new. On the other hand, damage to the pitch processing system would affect not only the processing of new pitch sequences but also the recogni-

tion of old ones. A useful analogy can be drawn here from the face-recognition literature which describes patients who are unable to recognize faces as familiar or unfamiliar, but who are still capable of distinguishing faces from nonfaces and performing a variety of perceptual matching tasks with them (e.g., DeRenzi, 1982). These patients are assumed to have damage to aspects of the face-recognition system, but to have structural analysis of faces intact. Conversely, there are also patients, thought to be impaired in structural analysis of faces, who are unable to perform facial matching tasks and also have difficulties distinguishing familiar from unfamiliar faces (e.g., Bodamer, 1947). A further neuropsychological dissociation that could arise is that the activation/retrieval of already well known tunes might be preserved with the ability to learn new tunes impaired. Such patients should be able to perform pitch processing and recognize old from new tunes, but be incapable of committing new melodic sequences to memory. As is often the case in psychology, these matters await empirical test, and we fully appreciate the dangers of prosecuting psychology in one area by analogy with another, but the key point we make here is that the two-systems view affords clear predictions.

If separation between access and activation of melodic entries looks likely although not yet fully established, we acknowledge that the evidence is even more mixed as to what is the actual division of labor between the pitch processing and melodic memory systems. Care is needed in allocating functions solely to access or lexicon, or to some close, but defined cooperation between the two. For example, interval processing and tonal functions may depend in part on long-term memory for either specific musical sequences or general tonal knowledge, presumably based on exposure to appropriate sequences within a genre. Hence, tasks investigating these issues may be better considered as mixed and involve a contribution by the lexicon to pitch processing itself. In this regard, it is interesting to note that Dowling and Bartlett (1981) found specific information regarding chromas or interval sizes to be more important in long-term memory tasks than in short-term memory tasks, with contour more relevant for short-term memory. For contour to be fully effective, however, it appeared that the melody had to be well learned and familiar (Dowling & Fujitani, 1971).

In this section, we have covered the major functions, and structural interrelationships with other components, of the pitch analysis subsystem. We now consider the rhythm analysis subsystem.

RHYTHM ANALYSIS SUBSYSTEM

Whereas pitch is concerned primarily with the sound aspects of the stimulus, rhythm is concerned primarily with its temporal characteristics, especially the ratio of stimulus and interstimulus durations and, we shall argue,

the representation of these as a timing sequence used to control rhythmic output. Having made this distinction, we are alert to the fact that we have previously implied that issues of serial ordering are crucial to the representation of pitch contours or trajectories. Such ordering would also seem to be crucial to capturing the correct sequence of a rhythmic structure and ultimately to correlate analyses of both pitch and rhythm. Thus, analyses of pitch and rhythm as a whole are likely to be weakly modular in the sense that both require access to systems sensitive to serial organization. Nevertheless we argue that there are modules specific to rhythmic processing that are not essential for pitch and vice versa (cf Large & Kolen, 1994). In the case of pitch, we have just discussed operations carried out by the pitch analysis subsystem. In the case of rhythm, there will be subsystems related to the timing and programming of motor representations. To develop this argument, we first examine the concept of rhythm in some detail, examine tasks that have been used to study it, and consider its neural basis.

Rhythm can be defined broadly or narrowly. By restricting its definition to music, it can be equated simply with the periodic accent and duration of notes. This has the advantage that it ties the concept to music, but the disadvantage that its links with other areas of psychological functioning governed by or specialized to deal with rhythm that may be excluded a priori. Too broad a definition, on the other hand, for example equating rhythm with any temporal variation in stimuli or responses, has the advantage that it does not prejudge the relationship between rhythm, in the musical sense, and other rhythmic manifestations, but may not be sufficiently precise to confine consideration to that set of rhythms governed by the same set of psychological principles. In practice, of course, it is impossible to be sure of the correct level of analysis. As with many other areas of psychological inquiry, it is likely that definitions will emerge and change as work progresses. Nevertheless, we adopt the following definition on the grounds that it appears to be broad enough to permit us to examine whether there are a common set of psychological principles underlying all rhythmic processing, yet, sufficiently precise to rule out trivial or more distant manifestations of temporal change such as breaking glass, or the temporal variations evident in, say, the pronunciation of single phonemes.

Thus we define rhythm as the temporally organized patterns resulting either from the initiation and termination of elements of a stimulus sequence within the same modality, or as the pattern resulting from discrete changes in one stimulus attribute (e.g., regular increases or decreases in loudness or brightness) of roughly equal magnitude, or as the pattern manifested by a set of regular responses. Rhythmic patterns varying in one perceptual attribute such as loudness or brightness need not covary in any other dimension such as pitch, color, pressure, or location. Typically a rhythmic sequence will consist of stimuli incorporating temporal discontinuities.

Expressed thus, the definition is broad enough to encompass stimulus sequences in, say, the visual or the tactile modality as well as audition (see for example Mavlov, 1980). The same rhythm may underpin a sequence of taps on a table, a sequence of auditory stimuli, or a sequence of light flashes, although whether or not these are ultimately captured by the same processing system is an empirical matter. For simplicity, the definition is phrased to exclude sequence switches from one modality to another, as with, say, a sequence such as three light flashes, followed by two auditory signals, followed by three light flashes, followed by two auditory signals, though it could be easily extended to accommodate these. The definition is also broad enough to include well-formed and ill-formed rhythmic sequences. Thus, we do not require that rhythmic stimuli necessarily have meta-qualities of beat, nor abide by musical rules captured by time signatures, although many such stimuli will do so in practice. Also, and crucially, rhythm can be exhibited by stimulus sequences that do not vary in other ways. This means that rhythmic information alone must be sufficient to specify a rhythmic sequence and to permit its analysis; although rhythmic information may in practice often covary with other types of information, such as pitch, it need not necessarily do so. Finally we include rhythms that can be perceived in human movements and responses, such as tapping, nodding, blinking, or limb flexing.

Despite the breadth of our definition, rhythmic stimuli in empirical research are typically more constrained. Most if not all rhythmic stimuli operate over a restricted time range in which elements of a rhythmic sequence are of the same durational order, in which elements typically last for comparable durations, and where the durations of "long" elements are fixed, small multiples of the durations of "short" segments. In practice, this time range is restricted by the processing limitations of human participants. The time envelope for subjective perceptions of rhythm appears to be between 115 and 1580 msec (Bolton, 1894), and studies of tempo suggest that around 100 beats per minute or around 600 msec per item is preferred and corresponds to the time for optimal temporal resolution (Fraisse, 1982). Many such stimuli also typically exhibit regularity in which rhythmic elements (stimuli and interstimulus intervals, ISIs) of a given sequence can be described as members of a restricted set of durations and repetitions in which a previously used rhythmic sequence may recur (see Large & Kolen, 1994, for further discussion of these qualities).

Before embarking on our consideration of studies of rhythmic processing, some preliminary points need to be made to put this section into context. Our reading of the literature suggests that many relevant studies of rhythm are to be found in the large literature on motor activity and the somewhat less extensive though still considerable literature on temporal perception and processing. These literatures have their own agendas, and,

although we hope to show they overlap with work on music perception and cognition, they include large bodies of work not strictly relevant to our account. Hence our review is necessarily selective. The literature dealing with the neurophysiology of motor activity is also exceedingly complex in that a wide variety of brain regions have been implicated in the production and control of movement sequences. The same is true of the timing literature. Within both of these broader domains, many issues still need to be resolved. In the case of rhythm in particular, we do not believe that a precise or fine-grained mapping is yet possible between the (many) cognitive processes needed to deal with rhythm and the neural systems that support them. Despite this, it is possible to identify the major functional aspects of cognitive systems involved in rhythm processing on the one hand, and the brain systems that seem capable of supporting such operations on the other, in sufficient detail to specify a model that, although it does not yet map onto a precise neural "wiring diagram" is nevertheless sufficiently specified to permit testable predictions to be drawn about its cognitive-neural components. To construct such a model, we have drawn attention explicitly to what appear to be various broad dissociations in studies of patients with brain injuries. As the data specific to these dissociations are still limited, our conclusions must be treated as provisional, but we offer our suggestions as more detailed alternatives to attempts to localize rhythm processing in one or the other cerebral hemisphere (Peretz, 1990). In discussing rhythm, we have purposely excluded issues of serial ordering, which we believe to be more relevant to the issue of *what next* and which we construe as the province of the pitch analysis system, rather than *when next,* which is properly the domain of rhythm (cf Large & Kolen, 1994). Nor have we yet considered the possible links between rhythmic and articulatory processes, although, as we shall argue later, these links will ultimately have to be addressed.

In the empirical literature, various motor tasks have been devised using such stimuli to examine the nature of rhythmic processing. These include tapping, auditory pacing, rhythm discrimination, and rhythm reproduction (e.g., Fraisse, 1982; Fries & Swihart, 1990; Halsband, Ito, Tanji, & Freund, 1993; Ivry & Keele, 1989; Keele, Pokorny, Corcos, & Ivry, 1985). Tapping tasks, without any stimulus present, have been used in a variety of ways, usually in the motor skills literature, both to study motor responding directly (Keele, Ivry, & Pokorny, 1987) and as a secondary task to interfere with other mental activity (Radil, Mates, Ilmberger, & Poppel, 1990; Radil, Mates, & Poppel, 1991). Auditory pacing can be thought of as a rhythmic tracking or shadowing task in which the participant is required to match or synchronize some motor response (typically tapping) to an external stimulus such as a visual or auditory rhythmic sequence (Fries & Swihart, 1990; Ivry & Keele, 1989). By contrast, rhythm discrimination, rhythm repro-

duction, and rhythm recognition all involve memory. Rhythm discrimination requires participants to detect the similarity or difference between two successively presented rhythmic sequences (Jones, Summerell, & Marshburn, 1987; Peretz & Kolinsky, 1993), rhythm reproduction requires participants to reproduce a previously heard rhythmic sequence (Ivry, Keele, & Diener, 1988), and rhythm recognition requires participants to perform a match between a current sequence and a previously presented one (Poulton, 1974). Such memory tasks permit additional variables, such as the length of the retention interval, or the nature of the interpolated material, to be manipulated, which auditory pacing or tapping alone do not.

A range of conditions have been identified that result in rhythm impairments of various sorts. These include Parkinson's disease (e.g., Benecke, Rothwell, Dick, Day, & Marsden, 1987; Crawford, Henderson, & Kennard, 1989; Jordan, Sagar, & Cooper, 1992), Huntington's chorea (e.g., Halsband, Hilperath, & Lange, 1990b; Hefter, Homberg, Lange, & Freund, 1987; Lasker, Zee, Hain, Folstein, & Singer, 1987), damage to the motor areas of the cortex (Dick, Benecke, Rothwell, Day, & Marsden, 1986; Freund & Hummelshein, 1984; 1985; Halsband, Ito, Tanji, & Freund, 1993), damage to the basal ganglia (e.g., Fries & Swihart, 1990; Laplane et al., 1989), and damage to the cerebellum (e.g., Arkhipova, Troshina, & Shabalov, 1991; Brown, Hefter, Mertens, & Freund, 1990; Inhoff & Rafal, 1990; Ivry et al., 1988; Stein, 1986).

Here we draw particular attention to two recent studies that, combined together, appear to establish a double dissociation between performance on direct auditory pacing and other tasks involving rhythm memory. Fries and Swihart (1990) reported the case of a previously musically skilled patient who, after damage of the right temporal areas and right basal ganglia, was seriously impaired in matching his tapping responses to a metronome beat, or tapping with the beat of marching band music, but who had no difficulty in perceiving rhythms or producing them from memory. The specificity of the patient's impairment to auditory pacing was further shown by his intact skill in tapping in response to rhythmic visual and tactile stimuli.

By contrast, Halsband et al. (1993) report studies of several patients with unilateral lesions in the lateral or medial premotor cortex (PMA) and supplementary motor area (SMA) who were profoundly impaired in the reproduction of rhythmic sequences from memory, but in the case of the SMA group were capable of auditory pacing (the PMA group had mixed etiology and included some patients with SMA lesions who were not tested on auditory pacing). In a further experiment, patients with SMA injuries were asked to retrieve a sequence of finger movements from memory or under visual guidance. Relative to controls, the patients with SMA injuries took much longer to initiate the finger sequence from memory with little or no difference under visual guidance.

An outline of two routes from rhythmic input to response can be hypothesized on the basis of these studies (see Figure 3). The figure illustrates what we shall refer to as the direct, or S-R route, between input and behavior, and a second, more circuitous route in which auditorily presented rhythms are processed and reproduced from memory (cf Halsband et al., 1993, for a similar conclusion). We further assume that both of these routes rely in the first instance on input into the auditory buffer, although other routes in, for example through visual and tactile modalities, are no doubt possible, we shall not consider them for now.

Examining each of these routes in more detail reveals further information about their cognitive and neurophysiological structure. Considering first the direct route, we note that the patient studied by Swihart and Fries sustained damage to the basal ganglia. The motor skills literature strongly implicates these structures in the control and output of movement (see, e.g., Connor & Abbs, 1990, for a review), and it is known that Parkinson's and Huntington's disease, both of which are known to affect the basal ganglia, result in disorders of the motor system including sequential and predictive voluntary movement (Stern, Mayeux, Rosen, & Ilson, 1983) and the acquisition of new motor skills (Sasaki & Gemba, 1981; Taylor, 1978). Moreover, the basal ganglia are thought to modulate movement sequences by influencing the selection of correct parameters needed for successful movement (Marsden, 1987). Caution is required here, however, in describing the direct route deficit, together with spared memory route, solely in terms of damage to the basal ganglia, because patients with Huntington's chorea in particular have been shown to be impaired on tasks involving rhythm memory (Halsband, Hilperath, & Lange, 1990b), patients with Parkinson's disease have been shown to be capable of using a predictive motor strategy of the sort presumably required for auditory pacing (Day, Dick, & Marsden, 1984), and auditory pacing may be impaired in patients with thalamic and upper midbrain abnormalities (case report cited in Fries & Swihart, 1990). Attributing impaired auditory pacing solely to damage of the basal ganglia is additionally risky given that Fries and Swihart's patient also sustained damage to the right temporal lobe on its lateral surface and exhibited other psychological impairments including disturbances of
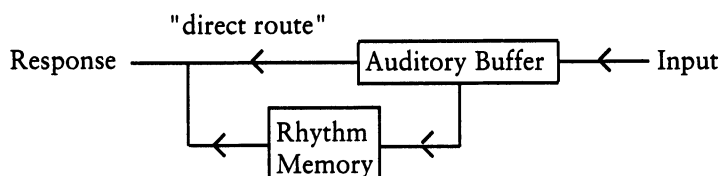

Fig. 3. Two routes in rhythmic processing.

pitch, time, and timbre discrimination and tonal memory. Thus their own conclusion that their patient was suffering from a central hearing disturbance may be safest at present.

Is the general distinction between pitch and rhythmic processing that we have been defending undermined by the concomitant impairment in pitch processing suffered by Swihart and Fries's patient? We think not. Despite the severity of the patient's impairments, (a) reproduction of rhythmic sequences from memory was unimpaired, suggesting that sufficiently detailed information about temporal sequencing of the auditory input was available for undamaged subsystems specialized for rhythm analysis, retention, and reproduction, and (b) tapping to visual sequences was unimpaired, implicating a problem specific to the auditory-motor route.

Turning now to the patients examined by Halsband et al. (1993), although both patients with lesions of the PMA and patients with lesions of the SMA showed disorders of rhythm memory, it can only be definitively concluded that patients with SMA lesions showed intact auditory pacing because their patients with PMA lesions were not tested in this way. Nevertheless, and in the light of the Swihart and Fries data, the suggestion by Halsband et al. that "a crucial distinction may be made between tasks in which the subject is asked to perform a given motor task where the model for performance is physically present at the time of performance, and those where the model is not present" (p. 262) is most probably correct. In line with this, they interpret their findings as implicating the SMA in "the preparation of internally remembered motor sequences" (p. 262). This view receives support from PET scans of persons asked to image but not to perform motor sequences in which activity in the SMA is detected (Roland, Larsen, Lassen, & Skinhoj, 1980), suggesting that motor plans can be activated and prepared but not executed if required (Heilman, Bowers, Valenstein, & Watson, 1987; Marteniuk, McKenzie, Jeannerod, Athenes, & Dugas, 1987).

As well as the basic distinction between the direct, stimulus response mapping of rhythmic sequences (auditory pacing), and the processing of rhythmic sequences from memory, finer distinctions can be drawn within rhythm memory tasks by considering a further aspect of the data from Halsband et al. (1993). Together with severely impaired skills of rhythm reproduction, their patients nevertheless showed intact rhythm discrimination. In the latter task patients were asked to check whether two successively presented rhythms were the same or different. We find it more useful to refer to this task as one involving rhythm recognition, albeit over a short time interval, rather than discrimination, because some record of the first rhythm must necessarily be formed and retained for comparison with the second. We submit that this recognition performance relies on some prior temporal analysis of previous rhythmic structure or on some simultaneous

comparison of the two rhythms in a subsystem sensitive to rhythmic temporal structure. This subsystem, presumably sufficient for the key activities involved in recognition, and undamaged in the patients studied by Halsband et al., is clearly not sufficient, although (as we shall argue) it is necessary for rhythm reproduction/recall, as the patients in the study by Halsband et al. were impaired in the latter but not the former.

To summarize, our analysis thus far suggests that rhythmic processing has three major aspects. First, rhythms can be copied directly through auditory pacing; second, they can be compared and matched (recognized); third, they can be recalled and reproduced. Double dissociations exist between the first activity on the one hand and the second and third activities on the other in that auditory pacing can be impaired without affecting rhythm memory and vice versa. In addition, a single dissociation has been demonstrated between rhythm recognition and rhythm reproduction with recognition spared when reproduction is impaired. To our knowledge, the reverse dissociation of impaired discrimination with spared reproduction of rhythms after recent or remote (preinjury) perception has not been demonstrated.

This pattern of spared and intact abilities suggests that there are key neural/cognitive subsystems uniquely involved in rhythm memory tasks that are not strictly required for auditory pacing (and vice versa), and additional subsystem(s), or a completely distinct subsystem, is needed for rhythm reproduction beyond the subsystems used for rhythm recognition. The division between the memory and pacing routes (Figure 3) captures only the first of these distinctions; to do justice to the recognition/reproduction distinction, further specification of the subprocesses in the rhythm memory is required. Here we offer a sketch of the general type and arrangement of processing components needed to perform rhythm memory tasks.

Based on the data from Halsband et al. (1993), we have just argued that there are two major subsystems involved in rhythm memory, differentially involved in rhythm recognition and reproduction. A critical question arising at this juncture is the one we have just alluded to, namely, to what extent do the tasks of rhythm recognition and reproduction rely at some point in their processing sequence on entirely separate processing components, or is their relationship such that subsystems for recognition are necessary but not sufficient for rhythmic reproduction? The reader can get the feel for these alternatives by examining Figure 4, in which two arrangements of components 1 and 2 of the rhythmic memory system are depicted. In Figure 4a, the two components are arranged such that reproduction and recognition rely on independent components dedicated to one task or the other, whereas in Figure 4b, components 1 and 2 are arranged sequentially and used differentially such that rhythm recognition (at least over the short retention intervals tested to date), can be accomplished using only compo-
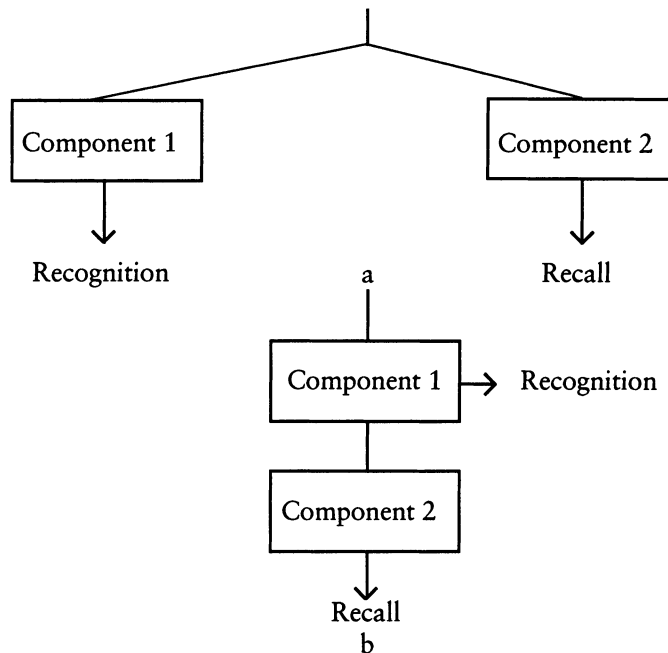
Fig. 4. Two possible arrangements of rhythm memory processing components: (a) complete independence of recognition and recall (production) and (b) rhythm recognition and production share a common processing component.

nent 1, while rhythm reproduction requires both components 1 and 2. According to Figure 4a, it should be possible to discover patients with the reverse dissociation to that reported by Halsband et al. Such individuals would show intact reproduction of rhythms, including newly acquired ones, but impaired recognition. The sequential arrangement of components (Figure 4b), on the other hand, implies that (1) whereas damage to component 2 will impair reproduction without affecting recognition, (2) damage to component 1 will affect both the recognition and reproduction of recently presented rhythmic sequences. The sequential arrangement also predicts that such patients, with damage to component 1, might still be capable of reproducing, from memory, rhythmic sequences learned before injury, while being incapable of reproducing newly presented sequences.

It is impossible to adjudicate between these alternatives on the basis of the data collected by Halsband et al. alone. Although seemingly implausible, and as yet undemonstrated, intact reproduction of recently presented sequences with impaired recognition cannot be ruled out at least as a theoretical possibility. There is, however, evidence consistent with damage to component 1 in a sequential arrangement. In a series of studies, Ivry and associates have provided good evidence for a common component needed for both the perception and production of timing tasks, damage to which

impairs both discrimination and production, and evidence for damage to
timing production but unimpaired perception (e.g., Inhoff & Rafal, 1990;
Ivry & Keele, 1989; Ivry et al., 1988). For example, Ivry and Keele con-
trasted different patient groups on a timing production and a timing per-
ception task. Participants included patients with cerebellar injury, patients
with Parkinson's disease, patients with cortical injury, and elderly and stu-
dent aged control subjects. Only the patients with cerebellar injury showed
impairments in the perception and the production tasks involving tempo-
ral processing, with the strong indication that the lateral region is impli-
cated in timing discrimination and generating short-term predictions of
movements. Patients with cortical injuries performed poorly on the pro-
duction task; as did cerebellar patients, they exhibited high variability in
tapping from memory to a standard compared with the elderly control
subjects, but were unimpaired on temporal perception. Patients with
Parkinson's disease performed comparatively well. They showed little im-
pairment in temporal perception and, although faster at tapping than the
elderly control subjects, had significantly less temporal variability when
tapping than the other patients with brain injuries. On the basis of these
data, Ivry and Keele argue that patients with cerebellar injuries had dam-
age to the common timing mechanism required for both tasks.

Of course, basing an argument for a shared mechanism in two tasks,
(recognition and reproduction) on a neuropsychological association such
as that observed in Ivry and Keele's patients with cerebellar injury may not
of itself be thought sufficient evidence that such a mechanism exists. There
are many reasons why an association might be discovered that have little to
do with functional identity or commonalities between tasks. However, there
is other converging evidence for a common timing mechanism shared by
perception and production in work conducted by Ivry and Hazeltine (1995)
on unimpaired participants. They have computed separate Weber func-
tions for the perception and production of temporal intervals in unimpaired
participants and find that when the two tasks both involve a single presen-
tation of the target interval per production or judgment, the slopes of the
functions are almost identical. Ivry and colleagues argue that this timing
mechanism is needed across a wide range of tasks, both motor and percep-
tual (Ivry & Hazeltine, 1995; Keele & Ivry, 1991).

We therefore adopt the sequential arrangement as the most likely, and
we refer to component 1 as the temporal analysis subsystem. This system is
required to process the incoming rhythmic pattern in terms of the relative
duration and number of stimuli and their interstimulus intervals. Such an
analysis is likely to be "on-line" and constrained by the absolute duration
of a rhythmic sequence and so, like the pitch subsystem, it will operate over
local rhythmic sequences rather than globally. Whether the timing mecha-
nism operates as a "processor" of temporal intervals (Ivry & Hazeltine,
1995), or as an internal oscillator that is "entrained" by or "resonates" to

external signals (Large & Kolen, 1994), or in some other way remains to be seen.

Component 2, assumed to be damaged in the patient studies by Halsband et al. and in the patients with cortical injury studied by Ivry and Keele, we hypothesize to have two functions. First it receives and retains the output of the temporal analysis subsystem. As such, it can be thought of as a more durable store for rhythmic sequences. Second, it is involved in the control and execution of rhythmic sequences from memory resulting in motor output such as tapping. Thus, to the extent that component 2 has the dual functions of storage and output control, we refer to it as the rhythm plan subsystem. Whether these storage and output functions turn out to be dissociable remains to be seen. The incorporation of these two components into the rhythm processing system is illustrated in Figure 5.

At present, we can only offer some suggestions about how a rhythm plan may be represented. On grounds of cognitive economy, its representation must be flexible enough to function in a variety of tasks ranging from the identification of rhythms as new or well known, to tapping, humming, or imaging rhythmic sequences. We assume the rhythm plan to be represented in a similar way to temporal components of a motor plan needed to govern an action sequence, that is, it is abstract in that it captures the essential, invariant aspects of the timing component of an action sequence, but it is not tied to any particular input or output system, specific set of movements, or even absolute timing. It is also complete, in that it specifies an entire rhythmic sequence that could turn out to be quite long, not merely
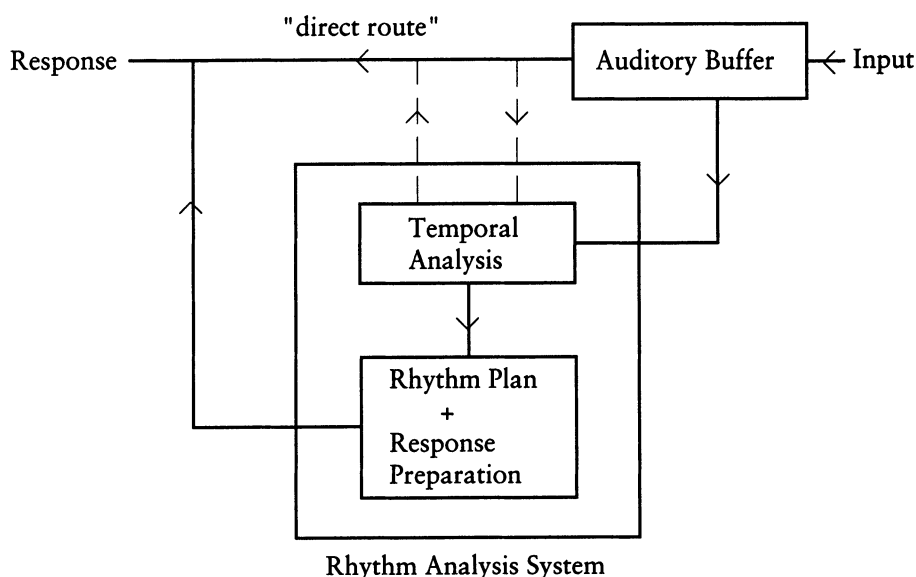

Fig. 5. An extended model of rhythmic processing.

its parts. Thus a key feature of the model is that the long-term representation of rhythmic sequences is as closely linked to the timing systems as it is to the auditory system. Finally, given the close relationships between the timing and motor systems, we suggest that the rhythm plan subsystem incorporates or is closely linked to a further subsystem for the execution, control, and ultimately the output of stored rhythm plans.

To summarize, the involvement of these rhythm memory subsystems in the major rhythmic tasks is as follows. The temporal analysis subsystem will be required for all tasks involving memory, namely recognition and reproduction. It may also be helpful, but not essential, for direct auditory pacing (dotted lines in Figure 5; Radil et al., 1990). In the case of recognition, particularly of short sequences retained over short intervals, temporal analysis will be sufficient for an effective match/mismatch to be made between the present and the previous stimuli. For recall, temporal analysis is necessary for the construction of the required rhythm plan, but insufficient for full plan storage and rhythm reproduction. Recall requires a prior temporal analysis, followed by its storage as a rhythm sequence, which can then function as the temporal component of a motor plan. The latter, in turn, will need to be selected and prepared for reproduction and its final output controlled by the execution subsystem.

As with pitch, to attempt a precise neurologic localization of these cognitive subsystems is somewhat premature given the current state of the motor literature (cf Kosslyn & Koenig, 1992), but to recapitulate some of our earlier observations, three general statements can be made: First, as we saw earlier, the cerebellum has been implicated in a range of tasks involving timing (Bloedel, 1992; Ivry & Keele, 1989; Ivry et al., 1988). Second, the motor cortices are thought to be implicated in the representation of motor plans (Eccles, 1982; Freund & Hummelshein, 1984, 1985; Halsband & Freund, 1990; Halsband et al., 1993; Rolls, 1983). Third, the distinction between the selection and preparation on the one hand and execution and control of action sequences has been made before (Goldberg, 1985; Jeannerod, 1981, 1984; Kosslyn & Koenig, 1992). Current evidence seems to suggest that the SMA plays an important role in the development of the intention to act and is active when a motor task is imagined as well as executed (Roland et al., 1980) Intriguing though these observations are, it would be unwise to equate uniquely the temporal control, rhythm plan, and execution subsystems with the cerebellum, the SMA, and the PMA, respectively. For one thing: the cerebellum is known to have output control as well as input timing functions (Bloedel, 1992; Ivry et al., 1988; Thach, Goodkin, & Keating, 1992); second, Ivry and Keele (1989) are careful to point out that although damage to the cerebellum may disrupt the timing circuit, the circuit itself is not contained wholly within the cerebellum; third, we do not equate the rhythm plan with a motor program in its entirety but

with the temporal component of such a program, and, in any case, other subcortical structures, especially the basal ganglia (Graybiel, 1990), have a role to play in motor programming and planning. Nevertheless, a prediction that does follow from the model is that although a unique mapping from cognition to neurology is not yet possible, damage to the brain areas mentioned is highly likely to have consequences at least for the three psychological functions under discussion: timing, plan representation, and execution. Further and more precise mapping of psychological functions onto neurologic systems must await further research.

MELODIC STRING MEMORY SUBSYSTEM

Everyday observation and empirical evidence support the existence of a melodic string memory system. Randomly selected subjects can hum tunes from verbal cues (e.g., "John Brown's Body"), and it is known that participants possess long-term memories for tunes (Bartlett & Snelus, 1980). The system is obviously required to recognize melodic sequences presented on separate occasions, to be alert to similar themes within a piece perhaps separated by several minutes, to reproduce tunes from memory, and to provide input for further evaluative and associative processes.

Beyond this, little is known in any detail about the underlying form of the representations used by the melodic memory system, its mode of operation, or its neurologic status but a number of its properties can be hypothesized:

1. Melodic memory representations must be sufficiently flexible to permit tune recognition despite changes in scale, timbre, or instruments. This implies that they instantiate aspects of tunes invariant across such transformations, namely, the unique combination of intervals, contour, and rhythms needed to specify the tune. In our model, therefore, we propose that the melodic string is composed of two elements: an entry in a tune lexicon, in the form of a *pitch string* (of the entire melody) constructed from the successive outputs of the pitch processing subsystem, and *the address of the associated rhythm plan* created by the rhythm memory system. Thus the melodic memory system combines aspects of the products of both the pitch and rhythmic subsystems (Figure 6).

   Like Peretz, we assume that access to the melodic memory system, and hence tune recognition, is primarily through an analysis of pitch.

2. Melodic memory representations can be flexibly used. The same melodic memory representation presumably drives a whistled, hummed, sung, or imaged tune as well as permitting the recogni-
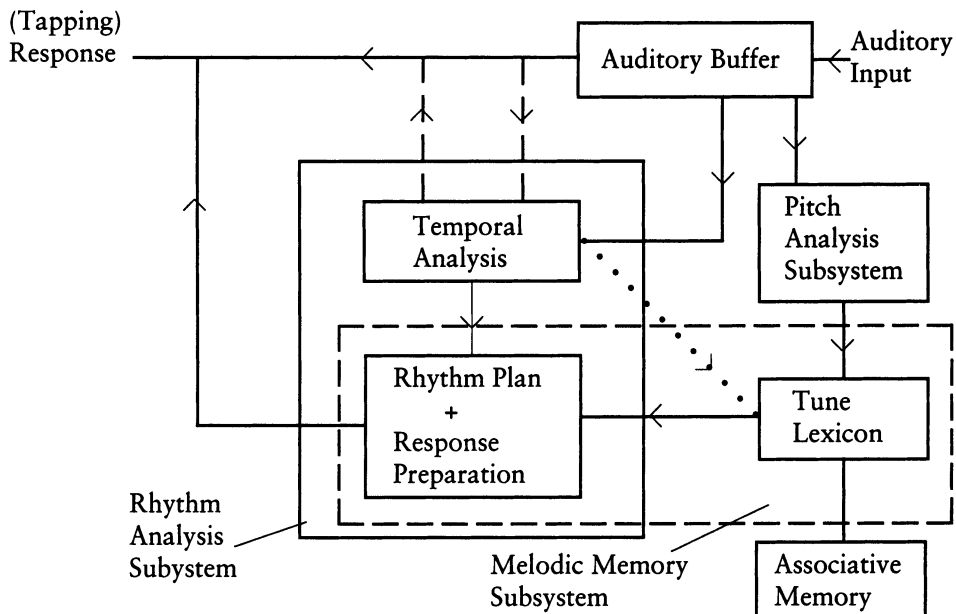
Fig. 6. Rhythm and pitch components and the melodic memory system.

tion of familiar tunes. Skilled performers on musical instruments, who "play by ear" may also be able to use a melodic representation to guide their playing without necessarily having sheet music in front of them or having been shown or having previously rehearsed the precise fingering, bowing, or other sequence of movements needed to produce the tune.

Tune reproduction, either overtly through singing or humming, or covertly using auditory imagery, starts with the activation of an entry in the lexicon. This can be cued in a variety of ways, for example, a snatch of the tune might be heard on the radio of a passing car, its title might be mentioned, or a remembered past experience associated with the tune might revive its entry. In all these cases, activation of the entry will simultaneously cue the contour string and, via its address, the associated rhythm plan.

3. Words are not a necessary part of a melodic memory representation. Anecdotally, one of the authors can whistle melodic passages from die Zauberflaute, but has no recollection of the German libretto. Objective evidence on this issue, however, is less straightforward. Thus, Samson and Zatorre (1991) have shown differential vulnerability of melodic and verbal codes to right and left hemisphere lesions, respectively, when words and tunes

are presented alone. However, they discovered that both lesions of the right hemisphere and lesions of the left hemisphere impaired recognition of tune and lyrics when these were presented together as sung text. These findings suggest that representations of lyrics and melody can be highly integrated in memory and jointly susceptible to damage. In a later study, however, lesions of the left temporal lobe had no effect on the performance on a lyric scanning task (Zatorre & Halpern, 1993). Zatorre and Halpern account for this discrepancy by noting that whereas the Samson and Zatorre study used new materials, the later study used highly overlearned songs. In such songs, they suggest, "the tune has become abstracted from the text to a large extent and can be accessed independently of it" (p. 229–230). Also, other patients have also been reported who cannot discern lyrics while recognizing tunes (Yaqub, Gascon, Alnosha, & Whitaker, 1988) or who cannot recognize tunes but who can recognize the words of familiar songs (Peretz & Kolinsky, 1993) suggesting that such dissociations do occur.

4. Strong links will exist between the melodic entry and associative memory for certain tunes. For example, tunes linked with rich environmental information such as films, strong autobiographical or emotional events in specific limited contexts, or subjected to extensive reflection, study, or analysis are likely to elicit elaborate and specific cognitive and emotional associations. On the other hand, new tunes, or tunes commonly heard in so wide a variety of contexts that they lose their episodic flavor are likely to have weaker associations, but in the case of familiar tunes at least, will still be readily classified into the appropriate semantic category, such as "Christmas" or "folk tune" (see Halpern, 1984, for evidence of the latter ability).

ASSOCIATIVE MEMORY

The necessity for associative memory has just been noted. Stimuli in any modality are capable of triggering vast amounts of associated information. In the case of musical stimuli, these will range from factual and nominal association to more personal, idiosyncratic and emotional ones. Suppose one is listening to Chopin's Nocturne No. 15 in F Minor. During this activity, a wide variety of cognitive and affective processes could be brought to bear, ranging from recall of the name and identity of the tune, awareness of the Polish origin of the composer, specific episodic memories of piano lessons when a child, emotional responses to the piece, and so on. Thus, like

Kosslyn and Koenig, we take this subsystem to include both semantic and episodic associations. More specifically, we view associative memory as relying heavily on declarative representations and as involving extensive amounts of relational processing and requiring the engagement of the hippocampal system for their formation and retrieval (see Cohen & Eichenbaum, 1993, for a detailed review).

The involvement of associative memory is the point at which both perception and imagery shade off into cognition. In general, therefore, the account we propose is better characterized as a General Systems Model rather than a narrow account of imagery—perceptual relations (cf Intons-Peterson, 1992). Such a General Systems Model proposes that imagery, perception, and cognition may recruit similar processes but to different extents depending on the demands of the task. Thus, certain imagery tasks may require largely perceptual judgments, whereas others will demand more cognitive activity. For example, compare the activities involved in answering the following: Which has the smoother fur, a terrier or a kitten? And now consider: Which would be the better playmate for a child? Both tasks will probably recruit imagery, in the first case largely tactile, in the second highly visual and possibly auditory, but the latter is likely to rely on more general knowledge and specific experience of cats, dogs, and children than the former. In each case, the processes induced will reflect the task demands, which is a view similar to Neisser's (1972) approach.

The locus of storage of associative memories is most likely to be in nonsensory cortical association areas, and, although we are not in a position to speculate about the precise neurologic bases of such general activity it seems likely that they will draw on information from a wide range of cortical areas and to make extensive use of linguistic categorization. However, it is known that certain specific amusias involving the identification of auditory stimuli may arise as a result of damage to auditory areas distinct from those involved in more basic perceptual analyses of audition. Vignolo (1982), for instance, has provided evidence showing that damage to the left temporal areas could produce impairment on a meaningless sounds identification test involving associations to auditory stimuli, unlike damage to the right temporal areas, which affected perception and discrimination as detected on a meaningless sounds discrimination test.

Toward a Theory of Auditory Imagery

Having now discussed the main components of the system required to perceive music, we finally consider, in general terms, how this system can give rise to auditory imagery for melodic sequences. We then end with a series of hypotheses about auditory imagery derived from the parent per-

ceptual theory. Most if not all of these require extensive testing, but it should be possible to do so using a combination of experimental methodologies, such as selective interference techniques, neuropsychological examinations, and direct brain imaging techniques. We have begun to examine several of these hypotheses in our own work but are keen for others to explore these too.

The orientation guiding all our hypotheses is that auditory imagery for musical sequences makes extensive use of the systems and structures used in music perception, especially the rhythm and pitch subsystems and the melodic memory system. To image an auditory sequence implies that information is fed back through the system and reanalyzed and, concomitantly, that preparatory motor activity takes place governing rhythm or articulation. There is already good evidence germane to this issue in studies by Zatorre, Halpern, and associates (e.g., Zatorre & Halpern, 1993; Zatorre et al., in press). Zatorre and Halpern (1993) studied the perception of and auditory imagery for songs in two groups of epileptic patients who had undergone excision of the left or right temporal lobe for severe epilepsy. In their perceptual task, participants heard a familiar song sung and simultaneously read its lyrics. On a given trial, they were required to judge whether the second of two capitalized lyrics was higher or lower in pitch than the first. The imagery task was identical to the perceptual one except that the song was generated using imagery rather than heard. Results clearly supported the idea that imagery and perception share common mechanisms. Although all patients found the imagery task harder than the perceptual task, the pattern of results for both tasks was similar in that the patients with damage of the right temporal lobe performed more poorly than those with left temporal damage. Zatorre and Halpern conclude that both auditory imagery and perception rely heavily on common neurologic mechanisms, particularly in the right temporal lobe. They also indicate that the impairment that they observe is not a simple pitch-processing deficit, but is likely to be specific to the auditory modality and to involve memory for melodic sequences. Support for and extension of these findings come from a cerebral blood flow study by Zatorre et al. (in press). They examined the neural correlates of auditory imagery and perception using the water bolus method of assessing regional cerebral blood flow in healthy participants who performed the same lyric task used by Zatorre and Halpern. Both imagery and perception resulted in similar patterns of neural activity with, as we stated earlier, increased bilateral activation in the secondary auditory cortices with some lateral asymmetry in the imagery task, left and right frontal lobes, left parietal lobes and the SMA, although the activation was considerably weaker in imagery than in perception. Comparisons between imagery and perception indicated increased activation when imaging in, among other areas, the frontal cortex and the hippocampus. Activation in

the temporal cortex was less strong and extensive in imagery compared with perception and tended to be concentrated away from primary auditory areas.

These results are particularly relevant for the current model. In general, they offer clear support for the proposition that perception and imagery tap similar, although not identical mechanisms, and support earlier work showing the importance of the secondary auditory cortices. However, a closer reading of their paper suggests a number of more specific relationships that now need to be explored.

First, and central to our own model, is our claim that rhythmic processes depend on areas not normally linked with audition, including parts of the motor system. Given this, it is intriguing to note that activation was observed by Zatorre and coworkers in the SMA. Although predicted by our model, we cannot be sure at present that this activity specifically reflects rhythmic processing because the task demands of the study by Zatorre et al. were such that participants were quite likely to have engaged in covert articulation. To what extent covert articulation, or "singing to oneself" is a psychological activity distinct or distinguishable from creating a rhythmically organized auditory image, or simply the only means by which a rhythmically organized auditory image can be constructed, remains to be seen. If singing to oneself is image creation, without residue, an interesting parallel is suggested between the sound and rhythmic dimensions of auditory imagery and the phonological and articulatory dimensions, respectively, of working memory (see also Gupta & MacWhinney, 1995, and Baddeley & Logie, 1992). Second, according to Zatorre et al., frontal structures are involved more when imaging than perceiving. Not surprisingly, imagery puts additional, constructive demands on participants, especially with tasks like lyric scanning, that necessitate central, attention-demanding cognitive operations in working memory likely to depend on these brain regions. These are aspects that go beyond our present model but need to be considered in due course. Third, little activation was discovered by Zatorre et al. in the primary auditory cortex when imaging. This contrasts with evidence from the visual modality, where visual imagery has been shown to elicit activation in the primary visual cortex (Kosslyn et al., 1993) and may reflect a basic distinction between auditory and visual imagery representation. In terms of the present model, the reconstructed representations of auditory imagery may not extend as far back as the auditory buffer, but may rely more on activity in the pitch, rhythmic, and melodic subsystems. In this sense, its representations could be said to be skewed away from immediate input toward more central and output processes.

These general parallels, and differences, between auditory imagery and perception encourage us to derive a number of more specific hypotheses from our current model:

HYPOTHESIS 1

Long-term (inactive) auditory images are stored as representations in the melodic memory system. These are precisely those representations that result from the prior perceptual analysis of musical sequences, and they consist of a melodic string plus the address of an associated rhythm plan.

HYPOTHESIS 2

Underlying representations in the melodic memory system can be activated and used to generate auditory images. There are a number of ways in which melodic memory representations could be cued. Two important ways are by concurrent activity in the melodic system, when for example, a fragment of the tune in question is heard, or through associative memory, after hearing the name of the tune, its composer, or other related information. It should therefore be possible to examine the structure of a melodic memory representation through priming effects in which the effects of prior cueing of, say, the name, lyric fragments, pitch sequences, or rhythmic fragments could be examined on processing times of later segments of lyric, pitch, or rhythm. Halpern's lyric scanning task can be thus thought of as a special case of a more general procedure in which previous song or melody components are compared against later components.

HYPOTHESIS 3

Auditory images, like their perceptual counterparts, of musical sequences have dissociable rhythm and pitch components.

This hypothesis has a number of corollaries, including the idea that it should be possible to interfere selectively with pitch and rhythmic components. Preliminary work in our laboratory suggests that this is possible. For instance, in a recent study, we tested the idea that concurrent tapping would interfere selectively with imaged rhythmic sequences by affecting either the encoding, maintenance, or deployment of a rhythm plan, but have little or no effect on imaged pitch. Participants judged whether either two notes, separated by either 5 or 10 sec, had the same pitch or whether two rhythmic sequences, separated by the same retention intervals, were identical. While performing these tasks, participants in the experimental conditions either tapped to a regular beat, tapped a previously learned irregular rhythm, or performed these activities covertly. Comparisons with a no-tapping control condition indicate that overt or covert tapping of irregular rhythms strongly interfered with the ability to retain and match imaged rhythmic sequences, but had no effect on the pitch task. It is worth emphasizing that preparatory motor activity, covert tapping, disrupted rhythm memory, providing further support for the notion that the two rely on common mecha-

nisms. A possible criticism of this study, however, is that pitch pair judg-
ments may be intrinsically easier to make than rhythmic sequence judg-
ments, as the former consist of matches between single as opposed to mul-
tiple stimulus elements. Hence, the interference effect may arise in the rhythm
conditions because of competition for limited general-purpose rather than
process-specific resources. To explore this further, we are now examining
the effects of tapping interference on multiple pitch sequences as well as
rhythm sequences.

HYPOTHESIS 4

The rhythmic component, or articulatory component, of an auditory
image cannot be activated without recruiting neural systems known to be
involved in motor activity, especially those involved in the planning of motor
sequences. Damage to these neural systems will result in specific impair-
ments in the rhythmic components of auditory imagery.

We have already discussed evidence from Zatorre et al. (in press) that
indicates activity particularly in the SMA when imaging melodies. The cor-
ollary of this is that brain injury resulting in damage to this and other areas
will have implications for the imaging of melodic sequences in general and
rhythmic processes in particular. We predict that damage to neurologic
mechanisms known to be involved in timing will affect both the recogni-
tion and recall of rhythm sequences, whereas areas such as the SMA known
to be involved in motor preparation will affect rhythm recall and rhythm
recognition carried out over intervals that necessitate the formation and
reactivation of previously stored rhythm plans, but will have little or no
effect on rhythm recognition conducted over short retention intervals. By
contrast, damage to the timing and motor mechanisms in question should
have no effect at all on simple pitch processing.

HYPOTHESIS 5

Short sequences, unassociated with known melodic representations, can
be "kept alive" for short durations through continued activity in the pitch
and rhythm subsystems. In this way, short-term auditory memories can be
maintained for novel strings. Presumably, with sufficient exposure or re-
hearsal, these will result in the formation of new melodic representations,
otherwise they will decay rapidly.

The extent to which any or all of these hypotheses are supported by the
data remains to be seen. What is clear and testable is the prediction of the
present model that auditory imagery and auditory perception will be found
to comprise dissociable subsystems organized around pitch and rhythm.
Also testable is the idea that auditory imaging is an activity in which out-

put processes, particularly those involved with rhythm, timing, and possibly articulation are used to control the replay of pitch sequences. These processes, it is suggested, will be as important as those permitting a reperception of the sound of the original stimulus. In short, auditory imagery might be more profitably construed as the playing or performing of tunes in one's head, rather than merely hearing the tunes play.[1]

References

Arkhipova, N. A., Troshina, E. M., & Shabalov, V. A. (1991). On the role of deep structures of the human brain in genesis of potentials connected with motion. *Zhurnal Vysshei Nervnoi Deyatel 'nosti, 41*(4), 655–662.

Assal, G. (1973). Aphasie de Wernicke sans amusie chez un pianiste. *Revue Neurologique, 129,* 251–255.

Baddeley, A. D. (1990). *Human memory: Theory and practice.* London, England: Lawrence Erlbaum Associates.

Baddeley, A. D., & Lieberman, K. (1980). Spatial working memory. In R. S. Nickerson (Ed.), *Attention and performance,* Vol. VIII. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Baddeley, A., & Logie, R. (1992). Auditory imagery and working memory. In Daniel Reisberg (Ed.), *Auditory imagery* (pp. 179–197). London: Erlbaum Associates.

Bartlett, J. C., & Snelus, P. (1980). Lifespan memory for popular songs. *American Journal of Psychology, 93,* 551–560.

Beech, J. R. (1978). *Image scanning: More problems for a pure propositional interpretation.* Paper presented at the B. P. S. Annual Conference, University of York.

Benecke, R., Rothwell, J. C., Dick, J. P. R., Day, B. L., Marsden, C. D. (1987). Disturbance of sequential movements in patients with Parkinson's disease. *Brain, 110,* 361–379.

Bever, T., & Chiarello, R. (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.

Bloedel, J. R. (1992). Functional heterogeneity with structural homogeneity: How does the cerebellum operate? *Behavioral and Brain Sciences, 15,* 666–678.

Bodamer, J. (1947). Die Prosop-Agnosie. *Archiv fur Psychiatrie und Nervenkrankheiten, 179,* 6–53.

Bolton, T. L. (1894). Rhythm. *American Journal of Psychology, 6,* 145–238.

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

Boltz, M. (1989b). Rhythm and "good endings": effects of temporal structure on tonality judgments. *Perception and Psychophysics, 46,* 9–17.

Boltz, M. (1989c). Time judgements of musical endings: Effects of expectancies on the "filled interval effect". *Perception and Psychophysics, 46,* 409–418.

Boltz, M., & Jones, M. R. (1986). Does rule recursion make melodies easier to reproduce? If not, what does? *Cognitive Psychology, 18,* 389–431.

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

Brown, S. H., Hefter, H., Mertens, M., & Freund, H. J. (1990). Disturbances in human arm movement trajectory due to mild cerebellar dysfunction. *Journal of Neurology, Neurosurgery and Psychiatry, 53*(4), 306–313.

Bruce, V., & Young, A. W. (1986). Understanding face recognition. *British Journal of Psychology, 77,* 305–327.

Brust, J. (1980). Music and language: Musical alexia and agraphia. *Brain, 103,* 367–392.

Castellano, M. A., Bharucha, J. J., & Krumhansl, C. L. (1984). Tonal hierarchies in the music of North India. *Journal of Experimental Psychology: General, 113,* 394–412.

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

Cohen, N. J., & Eichenbaum, H. (1993). *Memory, amnesia and the hippocampal system.* Cambridge, MA: MIT Press.

Connor, N. P., & Abbs, J. H. (1990). Sensorimotor contributions of the basal ganglia: recent advances. *Physical Therapy, 70*(12), 864–872.

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

Crawford, T. J., Henderson, L., & Kennard, C. (1989). Abnormalities of nonvisually-guided eye movements in Parkinson's disease. *Brain 112,* 1573–1586.

Crowder, R. G. (1989a). Imagery for musical timbre. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 472–478.

Crowder, R. G. (1989b April). *Imagery for musical timbre.* Paper presented at the Conference of Perception of Structure, New Haven, CT.

Day, B. L., Dick, J. P. R., & Marsden, C. D. (1984). Patients with Parkinson's disease can employ a predictive motor strategy. *Journal of Neurology, Neurosurgery and Psychiatry. 47,* 1299–1306.

De Renzi, E. (1982). *Disorders of space exploration and cognition.* Chichester: Wiley.

Deutsch, D. (1969). Music recognition. *Psychological Review, 76,* 300–307.

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

Deutsch, D. (1972). Octave generalization and tune recognition. *Perception & Psychophysics, 11,* 381–389.

Deutsch, D. (1973). Interference in memory between tones adjacent in the musical scale. *Journal of Experimental Psychology, 100,* 228–231.

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

Deutsch, D. & Feroe, J. (1981). The internal representation of pitch sequences in tonal music. *Psychological Review, 88,* 502–522.

Dick, J. P. R., Benecke, R., Rothwell, J. C., Day, B. L., & Marsden, C. D. (1986). Simple and complex movements in a patient with infarction of the right supplementary motor cortex. *Movement Disorders, 1,* 255–266.

Dorgeuille, C. (1966). *Introduction a l'etude des amusies.* Doctoral dissertation, Universite de la Sorbonne, Paris.

Dowling, W. J. (1973a). The perception of interleaved melodies. *Cognitive Psychology, 1973, 5,* 332–327.

Dowling, W. J. (1973b). Rhythmic groups and subjective chunks in memory for melodies. *Perception & Psychophysics, 14,* 37–40.

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

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

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

Eccles, J. C. (1982). The initiation of voluntary movements by the supplementary motor area. *Archiv fur Psychiatrie und Nervenkrankheiten, 231,* 423–441.

Engle, R. W., Cantor, J., & Turner, M. (1989). Modality effects: Do they fall on deaf ears? *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 41*(2-A), 273–292.

Eysenck, M. W., & Eysenck, M. C. (1979). Processing depth, elaboration of encoding, memory stores, and expended processing capacity. *Journal of Experimental Psychology: Human Learning & Memory, 5,* 472–484.

Farah, M. J. (1984). The neurological basis of mental imagery: A componential analysis. *Cognition, 18,* 245–271.

Farah, M. J. (1985). Psychophysical evidence for a shared representation medium for mental images and percepts. *Journal of Experimental Psychology, General, 114,* 91–103.

Farah, M. J., & Smith, A. (1983). Perceptual interference and facilitation with auditory imagery. *Perception & Psychophysics, 33*(5), 475–478.

Finke, R. L. (1980). Levels of equivalence in imagery and perception. *Psychological Review, 87,* 113–132.

Finke, M. J. (1985). Theories relating mental imagery to perception. *Psychological Bulletin, 98,* 236–259.

Finke, R. A. (1989). *Principles of mental imagery.* Cambridge, MA: MIT Press.

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

Frances, R. (1958). *La perception de la musique* [The perception of music]. Paris: J. Vrin.

Freund, H. J., & Hummelshein, H. (1984). Permotor cortex in man: evidence for innervation of proximal limb muscles. *Experimental Brain Research, 53,* 479–482.

Freund, H. J., & Hummelshein, H. (1985). Lesions of premotor cortex in man. *Brain, 108,* 697–733.

Fries, W., & Swihart, A. A. (1990). Disturbance of rhythm sense following right hemisphere damage. *Neuropsychologia, 28*(12), 1317–1323.

Geiselman, R. E., & Bjork, R. A. (1980). Primary versus secondary rehearsal in imagined voices: Differential effects on recognition. *Cognitive Psychology, 12,* 188–205.

Geiselman, R. E., & Glenny, J. (1977). Effects of imagining speaker's voices on the retention of words processed visually. *Memory & Cognition, 5,* 499–504.

Goldberg, G. (1985). Supplementary motor area structure and function: review and hypotheses. *Behavioral & Brain Sciences, 8,* 567–615.

Goldenberg, G., Podreka, I., & Steiner, M. (1990). The cerebral localization of visual imagery: evidence from emission computerized tomography of cerebral blood flow. In P. J. Hampson, D. F. Marks, & J. T. E. Richardson (Eds.), *Imagery: Current developments* (pp. 307–332). London: Routledge.

Graybiel. A. M. (1990). The basal ganglia and the initiation of movement. *Revue Neurologique (Paris), 146,* 10, 570–574.

Gromko, J. E. (1993). Perceptual differences between expert and novice music listeners: A multidimensional scaling analysis. *Psychology of Music, 21,* 34–47.

Gupta, P., MacWhinney, B. (1995). Is the articulatory loop articulatory or auditory? Reexamining the effects of concurrent articulation on immediate serial recall. *Journal of Memory and Language, 34,* 63–88.

Halpern, A. (1984). Organization in memory for familiar songs. *Journal of Experimental Psychology: Learning, Memory and Cognition, 10,* 496–512.

Halpern, A. (1988). Mental scanning in auditory imagery for songs. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*(3), 434–443.

Halpern, A. (1989). Memory for the absolute pitch of familiar songs. *Memory & Cognition, 17,* 572–581.

Halpern, A. (1992). Musical aspects of auditory imagery. In: D. Reisberg (Ed.), *Auditory imagery.* London: Lawrence Erlbaum Associates.

Halsband, U., & Freund, H. J. (1990). Premotor cortex and conditional motor learning in man. *Brain, 113,* 207–222.

Halsband, U., Hilperath, F., & Lange, H. W. (1990b). Presymptomatic higher motor disturbances in Huntington's disease. *Acta Neurologica Scandinavica Supplement ,128,* 34.

Halsband, U., Ito, J., Tanji, J., & Freund, H. J. (1993). The role of premotor cortex and the supplementary motor area in the temporal control of movement in man. *Brain, 116,* 243–266.

Hampson, P. J., Marks, D. F., & Richardson, J. T. E. (Eds.). (1990). *Imagery: Current developments.* London: Routledge.

Heffner, H. E., & Heffner, R. S. (1986). Effect of unilateral and bilateral auditory cortex lesions on the discrimination of vocalizations of Japanese macaques. *Journal of Neurophysiology, 56*(3), 683–701.

Heffner, H. E., & Heffner, R. S. (1990). Effect of bilateral auditory cortex lesions on sound localization in Japanese macaques. *Journal of Neurophysiology, 64(3),* 915–931.

Hefter, H., Homberg, V., Lange, H. W., & Freund, H. J. (1987). Impairment of rapid movement in Huntington's disease. *Brain, 110, 585–612.*

Heilman, K. M., Bowers, D., Valenstein, E., & Watson, E. (1987). Hemispace and hemispatial neglect. In M. Jeannerod (Ed.), *Neurophysiological and neuropsychological aspects of spatial neglect* (pp. 115–150). North Holland: Amsterdam.

Hubbard, T. L., & Stoeckig, K. (1992). The representation of pitch in musical images. In D. Reisberg (Ed.), *Auditory imagery* (pp. 199–235). London: Erlbaum Associates.

Inhoff, A. W., & Rafal, R. (1990). Cerebellar structures and the programming of movement sequences. *Behavioral-Neurology, 3*(2). 87–97.

Intons-Peterson, M. J. (1980). The role of loudness in auditory imagery. *Memory & Cognition, 8,* 385–393.

Intons-Peterson, M. J. (1992). Components of auditory imagery. In D. Reisberg (Ed.), *Auditory imagery* (pp. 45–71). London: Erlbaum Associates.

Intons-Peterson, M. J., Russell, W., & Dressel, S. (1992). The role of pitch in auditory imagery. *Journal of Experimental Psychology, Human Perception & Performance, 18*(1), 233–240.

Ivry, R. B., & Hazeltine, R. E. (1995). The perception and production of temporal intervals across a range of durations: Evidence for a common timing mechanism. *Journal of Experimental Psychology: Human Perception and Performance, 21,* 3–18.

Ivry, R. B., & Keele, S. W. (1989). Timing functions of the cerebellum. *Journal of Cognitive Neuroscience. 1,* 136–152.

Ivry, R. B., Keele, S. W., & Diener, H. C. (1988). Dissociation of the lateral and medial cerebellum in movement timing and movement execution. *Experimental Brain Research, 73,* 167–180.

Jeannerod, M. (1981). Intersegmental coordination during reaching at natural visual objects. In J. Long & A. Baddeley (Eds.), *Attention and performance* (pp. 153–168). Hillsdale, NJ: Erlbaum.

Jeannerod, M. (1984). The timing of natural prehension movements. *Journal of Motor Behavior, 16,* 235–254.

Johnson, M. K., Foley, M. A., & Leach, K. (1988). The consequences for memory of imagining in another person's voice. *Memory & Cognition, 16,* 337–342.

Johnston, W. A., & Heinz, S. P. (1978). Flexibility and capacity demands of attention. *Journal of Experimental Psychology: General, 107,* 420–435.

Jones, M. R. (1974). Cognitive representation of serial patterns. In B. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition.* Potomac, MD: Erlbaum.

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

Jones, M. R., Boltz, M., & Kidd, G. (1982). Controlled attending as a function of melodic and temporal context. *Perception & Psychophysics, 32,* 221–218.

Jones, M. R, Summerell, L., & Marshburn, E. (1987). Recognizing melodies: a dynamic interpretation. *Quarterly Journal of Experimental Psychology, 39a,* 89–1121.

Jordan, N., Sagar, H. J., & Cooper, J. A. (1992). Cognitive components of reaction time in Parkinson's disease. *Journal of Neurology, Neurosurgery, & Psychiatry, 55,* 658–664.

Keele, S., & Ivry, R. (1991). Does the cerebellum provide a common computation for diverse tasks: a timing hypothesis. In A. Diamond (Ed.), *Developmental and neural basis of higher cognitive function* (pp. 179211). New York: Annals New York Academy of Sciences.

Keele, S., Ivry, R., & Pokorny, R. (1987). Force control and its relation to timing. *Journal of Motor Behaviour, 19,* 96–114.

Keele, S., Pokorny, R., Corcos, D., & Ivry, R. (1985). Do perception and motor production share common timing mechanisms: A correlational analysis. *Acta Psychologia, 60,* 173–191.

Kosslyn, S. M. (1975). Information representation in visual images. *Cognitive Psychology, 7,* 341–370.

Kosslyn, S. M. (1980). *Image and mind.* Cambridge, MA: Harvard University Press.

Kosslyn, S. M., Alpert, N., Thompson, W. L., Maljkovic, V., Weise, S. B., Chabris, C. F., Hamilton, S. E., Rauch, S. L., & Buonanno, F. S. (1993). Visual mental imagery activates topographically organized visual cortex: PET investigations. *Journal of Cognitive Neuroscience, 5,* 263–287.

Kosslyn, S. M., Ball, T. M., & Reiser, B. A. (1978). Visual images preserve metric spatial information: Evidence from studies of image scanning. *Journal of Experimental Psychology: Human Perception & Performance, 4,* 47–60.

Kosslyn, S. M., & Koenig, O. (1992). *Wet mind: The new cognitive neuroscience.* New York: Macmillan Inc.

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

Laplane, D., Levasseur, M., Pillon, B., Dubois, B., Baulac, M., Mazoyer, B., Tran-Dinh, S., Sette, G., Danze, F. & Baron, J. C. (1989). Obsessive-compulsive and other behavioural changes with bilateral basal ganglia lesions: A neuropsychological, magnetic resonance imaging and positron tomography study. *Brain, 112,* 699–725.

Large, E. W., & Kolen, J. F. (1994). Resonance and the perception of musical meter. *Connection Science, 6*(2,3), 177–207.

Lasker, A. G., Zee, D. S., Hain, T. C., Folstein, S. E., & Singer, H. S. (1987). Saccades in Huntington's disease: initiation defects and distractibility. *Neurology 37,* 364–370.

Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music.* Cambridge, MA: MIT Press.

Logie, R. H., & Baddeley, A. D. (1990). Imagery and working memory. In P. J. Hampson., D. E. Marks., & J. T. E. Richardson (Eds.), *Imagery: Current developments* (pp. 103–128). London: Routledge.

Lu, Z. L., Williamson, S. J., & Kaufman, L. (1992). Behavioral lifetime of human auditory sensory memory predicted by physiological measures. *Science, 258*(5088), 1668–1670.

Marks, D. F. (1990). On the relationship between body and mind. In P. J. Hampson, D. F. Marks, & J. T. E. Richardson (Eds.), *Imagery: Current developments* (pp. 1–38). London: Routledge.

Marsden, C. D. (1987). What do the basal ganglia tell premotor cortical areas? In G. Bock, M. O'Connor & J. Marsh (Eds.), *Motor areas of the cerebral cortex.* Ciba Foundation Symposium 132, (pp. 282–300). Chichester: John Wiley.

Marteniuk, R. G., Mackenzie, C. L., Jeannerod, M., Athenes, S., & Dugas, C. (1987). Constraints on human arm movement trajectories. *Canadian Journal of Psychology, 41,* 365–378.

Mavlov, L. (1980). Amusia due to rhythm agnosia in a musician with left hemisphere damage: a non auditory supramodal defect. *Cortex, 16,* 321–328.

McClelland, J. L., & Rumelhart, D. E. (Eds.). (1986). *Parallel distributed processing: Explorations in the microstructure of cognition: Vol. 2. Psychological and biological models.* Cambridge, MA: M. I. T. Press

Mendez, M. F., & Geehan, G. R. (1988). Cortical auditory disorders: Clinical and psychoacoustic features. *Journal of Neurology, Neurosurgery, and Psychiatry, 51*(1), 1–9.

Miller, G. A. (1967). *The psychology of communication.* Baltimore: Penguin Books.

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

Monahan, C. B., & Carterette, E. C. (1985). Pitch and duration as determinants of musical space. *Music Perception, 3,* 1–32.

Monahan, C., Kendall, R., & Carterette, E. (1987). The effect of melodic and temporal contour on recognition memory for pitch change. *Perception & Psychophysics, 41,* 576–600.

Morris, P. E., & Hampson, P. J. (1983). *Imagery and consciousness.* London: Routledge.

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

Nairne, J. S., & Pusen, C. (1984). Serial recall of imagined voices. *Journal of Verbal Learning and Verbal Behavior, 23,* 331–342.

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

Neisser, U. (1972). Changing conceptions of imagery. In P. W. Sheehan (Ed.), *The function and nature of imagery.* London: Academic Press.

Neisser, U. (1976). *Cognition and reality.* San Francisco: W. H. Freeman.

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

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

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

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

Penfield, W. (1967). *The excitable cortex in conscious man.* Liverpool: University Press.

Penney, C. G. (1989). Modality effects and the structure of short-term verbal memory. *Memory & Cognition, 17,* 398–422.

Peretz, I. (1987). Shifting ear-asymmetry in melody comparison through transposition. *Cortex, 23,* 317–323.

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

Peretz, I. (1993). Auditory agnosia: a functional analysis. In S. McAdams & E. Bigand (Eds.), *Thinking in sound: The cognitive psychology of human audition* (pp. 199–230). Oxford: Clarendon Press.

Peretz, I., & Babai, M. (1992). The role of contour and intervals in the recognition of melody parts: evidence from cerebral asymmetries in musicians. *Neuropsychologia, 30,* 277–292.

Peretz, I., & Kolinsky, R. (1993). Boundaries of separability between melody and rhythm in music discrimination: a neuropsychological perspective. *Quarterly Journal of Experimental Psychology, 46A(2)* 301–325.

Peretz, I., & Morais, J. (1987). Analytic processing in the classification of melodies as same or different. *Neuropsychologia, 25,* 645–652.

Peretz, I., Morais, J., & Bertelson, P. (1987). Shifting ear differences in melody recognition through strategy inducement. *Brain & Cognition, 6,* 202–215.

Petrides, M., Alivisatos, B., Meyer, M., & Evans, A. C. (1993). Functional activation of the human frontal cortex during the performance of verbal working memory tasks. *Proceedings of the National Academy of Sciences USA, 90,* 878–882.

Pitt, M. A., & Crowder, R. G. (1992). The role of spectral and dynamic cues in imagery for musical timbre. *Journal of Experimental Psychology, Human Perception & Performance, 18,(3),* 728–738.

Poulton, E. C. (1974) *Tracking skill and manual control.* New York: Academic Press.

Radil, T., Mates, J., Ilmberger, J., & Poppel, E. (1990). Stimulus anticipation in following rhythmic acoustical patterns in tapping. *Experientia, 46,* 762–763.

Radil, T., Mates, J., & Poppel, E. (1991). Rhythm extrapolation in tapping. *Homeostasis, 33(4),* 222.

Riddoch, M. J. (1990). Loss of imagery: A generation deficit. *Cognitive Neuropsychology, 7(4),* 249–273.

Roland, P. E., Larsen, B., Lassen, N. A., & Skinhoj, E. (1980). Supplementary motor area and other cortical areas in organization of voluntary movements in man. *Journal of Neurophysiology, 43,* 118–136.

Rolls, E. T. (1983). The initiations of movements. *Experimental Brain Research, 7*(Suppl.), 97–113.

Samson, S., & Zatorre, R. J. (1988). Discrimination of melodic and harmonic stimuli after unilateral cerebral excisions. *Brain & Cognition, 7,* 348–360.

Samson, S., & Zatorre, R. J. (1991). Recognition memory for text and melody of songs after unilateral temporal lobe lesion: Evidence for dual encoding. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 793–804.

Samson, S., & Zatorre, R. J. (1994). Contribution of the right temporal lobe to musical timbre discrimination. *Neuropsychologia, 32,* 231–240.

Sasaki, K., & Gemba, H. (1981). Cortical field potentials preceding self-paced and visually initiated hand movements in one and the same monkey and influences of cerebellar hemispherectomy upon the potentials. *Neuroscience Letters, 25,* 287–292.

Segal, S. J., & Fusella, V. (1970). Influence of imaged pictures and sounds on detection of visual and auditory signals. *Journal of Experimental Psychology, 83,* 458–464.

Serafine, M. L., Crowder, R. G., & Repp, B. H. (1984). Integration of melody and text in memory for songs. *Cognition, 16,* 285–303.

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

Shepard, R. N. (1982b). Structural representations of musical pitch. In D. Deutsch (Ed.), *Psychology of music* (pp. 343–390). New York: Academic Press.

Stein, J. F. (1986). Role of the cerebellum in the visual guidance of movement. *Nature, 323*(6085), 217–221.

Stern, Y., Mayeux, R., Rosen, J., & Ilson, J. (1983). Perceptual motor dysfunction in Parkinson's disease: A deficit in sequential and predictive voluntary movement. *Journal of Neurology, Neurosurgery and Psychiatry, 46,* 145–151.

Taylor, M. J. (1978). Bereitschaftspotential during the acquisition of a skilled motor task. *Electroencephalography and Clinical Neurophysiology, 45,* 568–576.

Thach, W. T., Goodkin, H. P., & Keating, J. G. (1992). The cerebellum and the adaptive coordination of movement. *Annual Review of Neuroscience, 15,* 403–442.

Tramo, M., Bharucha, J., & Musiek, F. (1990). Music perception and cognition following bilateral lesions of auditory cortex. *Journal of Cognitive Neuroscience, 2,* 195–212.

Treisman, A. M. (1964). Verbal cues, language, and meaning in selective attention. *American Journal of Psychology, 77,* 206–219.

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

Vignolo, L. (1982). Auditory agnosia. *Philosophical Transactions of the Royal Society of London, 298,* 49–57.

Weber, W. J., & Brown, S. I. (1986). Musical imagery. *Music Perception, 3,* 411–426.

Williams, D. B. (1982). Auditory cognition: A study of the similarities in memory processing for music tones and spoken words. *Bulletin of the Council for Research in Music Education, 71*(Summer), 30–44.

Yaqub, B., Gascon, G., Alnosha, M., & Whitaker, H. (1988). Pure word deafness (acquired verbal auditory agnosia) in an Arabic speaking patient. *Brain, 111,* 457–466.

Zatorre, R. J. (1985). Discrimination and recognition of tonal melodies after unilateral cerebral excisions. *Neuropsychologia, 23,* 31–41.

Zatorre, R. J., Evans, A. C., & Meyer, E. (1994). Neural mechanisms underlying melodic perception and memory for pitch. *Journal of Neuroscience, 14,* 1908–1919.

Zatorre, R. J., & Halpern, A. R. (1993). Effects of unilateral temporal-lobe excision on perception and imagery of songs. *Neuropsychologia, 31,* 221–232.

Zatorre, R. J., Halpern, A. R., Perry, D. W., Meyer, E., & Evans, A. C. (in press). Hearing in the mind's ear: a PET investigation of musical imagery and perception. *Journal of Cognitive Neuroscience.*

Zatorre, R. J., & Samson, S. (1991). Role of the right temporal cortex in retention of pitch in auditory short-term memory. *Brain, 114,* 2403–2417.