Perception of Structure in Auditory Patterns

Roisin L Ash

Thesis submitted to the University of Stirling in partial
fulfilment of the requirements for the degree of
Doctor of Philosophy

Acknowledgements

I would like to thank Roger Watt for his never-ending confidence and enthusiasm towards my work and for overall supervision and guidance throughout my PhD. I am eternally grateful to Bruce, Bob and Colin, the three knights in shining white coats, for coming to the rescue when equipment failed. Thanks must also go to: everyone who participated in my experiments; Essex Boy, Will, Trish & Harry for having been there; the Chalton Road Crewe and the Edinburgh mob for a great time outside work and for keeping me in touch with the 'real world'; Fiona for those amazing cooking and baking extravaganzas; Lawrence for pointing out that I shouldn't really be eating portions which are bigger than the size of my head; big Geordie for his wardrobe stunt and, of course, John for showing me the ultimate kitchen sin. Finally, a really massive thank you must go to Georgette, Ishbel and Donald for being Georgette, Ishbel and Donald, respectively, and for keeping me 'sane'.

Contents

Chapter 1 Introduction

Chapter 2 Segmentation and Click Localisation

Experiment I Segmentation

Experiment II Click localisation

Condition 1 Click localisation in Gaelic melodies

Chapter 3 Rhythm Recall

Experiment I Rhythm recall of unstructured versus structured pitch sequences

Experiment II Rhythm recall of Gaelic melodies

Condition I Effect of pitch and rhythm alterations on rhythm recall

Chapter 4 Beat Perception

Chapter 5 Perception of Musical Intervals and Melodies

Experiment I Perception of Musical intervals

Experiment II Perception of content in Gaelic melodies

Experiment III Perception of content in structurally altered melodies

Chapter 6 Discussion

Abstract

The present research utilised five tasks to investigate non-musicians' perception of phrase, rhythm, pitch and beat structure in unaccompanied Gaelic melodies and musical sequences.

Perception of phrase structure was examined using: i) a segmentation task in which listeners segmented Gaelic melodies into a series of meaningful units and ii) a novel click localisation task whereby listeners indicated where they perceived a superimposed click in the melody had occurred. Listeners consistently segmented the melodies into units of 2.4 - 5.4 seconds. Clicks which were positioned before and after perceived boundaries (identified by segmentation) were perceptually migrated towards the boundary. These results suggest that listeners perceptually differentiate between phrasal groups in melodies (See Sloboda & Gregory, 1980; Stoffer, 1985, for similar results with musicians).

Short term memory for rhythmic structure was examined using rhythm recall of computer generated sequences and Gaelic melodies. Computer generated rhythms with small tonal pitch intervals (1 - 4 semitones) were easier to recall than large atonal intervals (predominantly greater than 4 semitones). Recall of Gaelic melodies, containing repetitive rhythmic units, was better than recall of computer sequences. Pitch reversal of Gaelic melodies did not effect recall.

Beat-tapping with three Gaelic melodies revealed that the majority of listeners established the underlying beat 1.5 - 3 seconds (5 - 6 notes) after the start of the melodies.

Perception of meaning and content in two note melodic intervals and three Gaelic melodies was examined using an adjective pair two-alternative forced choice task. Responses to musical intervals showed evidence of perceptual similarity based mainly on interval size. Perceived information content in the melodies increased significantly by the fourth note.

The results suggest that the amounts of Gaelic melody which are: i) required to establish an underlying beat, ii) remembered after one hearing, and iii) perceptually grouped into a meaningful unit, include the unit of melody which is necessary to establish a basic meaning.

Chapter 1 Introduction

Background

Almost all of us listen to music of some description. The activity of listening to music is a universal cultural experience and, as such, music can be used to broaden our understanding of the human mind (Dowling & Harwood, 1986). Indeed, it is argued that music is one of the most perceptual and cognitive of all experiences (Tighe & Dowling, 1993). Over the past twenty or so years there have been significant increases in the empirical study of music by psychologists and musicians (Deutsch, 1982a; Aiello, 1994). The number of psychologists involved in the music perception field has increased dramatically over the last decade and is still increasing (Jones & Holleran, 1992). A wide variety of techniques have been employed to examine various aspects of music perception including neuropsychological research (Peretz, 1993), aesthetics (Crozier, 1974, artificial intelligence (Balaban, Ebcioglu & Laske, 1992), computational modelling (Desain & Honing, 1997), theoretical work on musical comprehension (Lerdahl & Jackendoff, 1983) and empirical psychology techniques such as discrimination of temporal patterns (Sternberg & Knoll, 1994) and musical intervals (Burns & Ward, 1978) and tasks such as recognition (Cuddy, Cohen & Mewhort, 1981) and immediate recall (Deutsch, 1980) of musical sequences.

Empirical studies have been directed largely to the study of the processes involved in music listening as opposed to music performance or composition (Sloboda, 1988). Understanding the cognitive element of music perception has appealed to many researchers. There are a number of prominent texts which examine the cognitive basis of general audition (McAdams & Bigand, 1993), musical pitch (Krumhansl, 1990) and musical communication (Jones & Holleran, 1992). It has been argued that cognitive psychology has focused mainly on speech perception, failing to address the issues surrounding our understanding of the ways in which music may reach its communicative goal (Jones & Holleran, 1992). There is a tendency for psychologists carrying out research in the music perception field to disseminate their knowledge to other professional psychologists with few interdisciplinary communications between psychologists and musicians (Jones & Holleran, 1992). They point out that there are often difficulties of widely differing terminology and theoretical perspectives which exist in a field of such diversity as music perception.

Musical structure

Despite the differences between research backgrounds and empirical approaches and techniques, there is one thing which all researchers agree upon: music has structure (Dowling & Harwood, 1986). It is the structure of music which is essential to the listeners' understanding. Musical structure commonly refers to the basic elements which make up the music, such as the pitch and rhythm patterns. The structures dealt with in greater detail in the present thesis are beat, rhythm, pitch and phrase structure. Each of these will be described briefly below.

1. Beat

The beat is the regular pulse of the music (Sloboda, 1985). We have all experienced reacting to the beat of music. A range of bodily movements such as foot and hand tapping, head nodding and walking, can be induced and controlled by the underlying beat of the music. Listeners can be unaware of the ways in which the beat of the music is controlling or contributing to their behaviour. These kinds of universal responses have led some researchers to the conclusion that beat extraction is one of the processes which is fundamental to our musical understanding (Drake, Penel, Bigand & Stefan, 1997).

2. Metre

The way in which most of the music we listen to invites us to tap our foot or follow a beat implies the presence of a metrical organisation (Bigand, 1993; Jones & Yee, 1993). Metre imposes an accent structure on the underlying beats of the music, resulting in a periodic alternation of strong and weak accents (Palmer & Krumhansl, 1990; Sloboda, 1985). Music is likely to be metrical in that it is arranged around regular beats which are grouped into bars or measures with a major accent on the first beat. The simplest type of metrical structure has a stressed beat on every second, third or fourth (or some multiple of these). Metre is the main periodic beat structure in the music which is shown by bar lines in the score (Sloboda, 1985; Parncutt, 1987).

3. Rhythm

Rhythmic structure refers to the temporal relationships between events such as notes and rests in the music. More specifically, rhythm refers to the pattern of successive durations in a sequence (Parncutt, 1987; Jones, 1987). Rhythm describes the temporal organisation of a melody and, as such, may directly affect the units or segments which are perceived in a given melody. It is necessary to perceive the underlying beat when attempting to understand the rhythmic structure in a melody (Longuet-Higgins, 1976). Indeed, the use of a beat framework in perceiving rhythmical structures is almost universal (Dowling & Harwood, 1986).

4a. Pitch

Pitch is commonly used to refer to how high or low a given note is. Along with rhythm, pitch is considered as one of the most basic psychological dimensions of music (Dowling & Harwood, 1986). Most of the music we listen to is constructed around a musical scale, commonly the diatonic scale, for example, the C Major scale which is made up of all the white notes on the piano. Musical scales demonstrate some of the pitch constraints which commonly occur in the music we listen to. The tonic is most often heard at the beginning and end of a piece of music. The pattern of successive pitch intervals which make up the melodies of much of the music we listen to are related to each other in meaningful ways which are constrained by the use of notes from a particular scale. It is the set of relationships between different pitches which create pitch interval structure.

4b Pitch contour

One way of describing the pattern of a sequence of pitches is to consider the pitch contour, or the series of pitch movements up and down between adjacent notes in a melody. The sequence C_4 - E_4 - G_4 - A_4 - F_4 would have a contour of +4 +3 +2 -4 which characterises the direction and size of adjacent pitch intervals in semitone steps.

5. Phrase

A phrase is described as a relatively complete musical utterance, usually with a clear beginning and with a sense of conclusion or cadence (Brandt, Corra, Christ, Delaine & Windold, 1980). A phrase in music can be used to denote a unit of information or meaning (Palmer & Krumhansl, 1987b; Tan, Aiello & Bever, 1981) which is made up of the combination of several smaller units or motives. Motives are described as the most basic fragments of a melody and consist of at least one characteristic musical interval and rhythm (Boyden, 1971). Tones in music combine to create rhythmic and melodic motives which in turn combine to create phrases and longer sections of musical pieces.

The role of memory

The term *music* describes a vast quantity of different kinds of patterned sound and naturally the structure is not the same in different kinds of music (Jones, 1987). In order that music achieves its communicative function, these kinds of musical structures need to be perceived by the listener, although not explicitly. The perception of metre, rhythm, pitch and phrase structure will be discussed in the next section. The very act of listening requires that the listener processes structured information over time (Palmer & Krumhansl, 1990). In order to understand present occurrences in a melody, the listener must have an idea of what has gone before (Dowling & Harwood, 1986). The perception and cognition of music obviously requires some form of memory component. When faced with a piece of music, listeners do not hear one big chunk of sound. Nor is it likely that they hear, for example, 167 isolated notes. Listeners engage in some kind of perceptual grouping process during music listening.

Short term memory

It is likely that subjective grouping of auditory information is related to the listener's perceptual processing capacities. Attention and memory are limited in the amount of incoming information with which they can deal during the course of listening to an auditory sequence (Miller, 1994; Fraisse, 1982). Grouping the stimulus into smaller subgroups allows listeners a better chance of handling and organising the incoming information into a series of meaningful and related units within which detailed processing can occur (Swain, 1986; Michon, 1985). Studies which examine short term memory (STM) limitations have attempted to discover the limit of this memory capacity. Verbal STM refers to the memory capacity which is used to remember, for example, a new phone number for the duration it takes you to find a pen, sometimes referred to as working memory (Baddeley & Hitch, 1974). The span of STM is typically examined using some form of digit span technique in which a list of digits is presented to participants and they are asked to recall the list. The number of digits which participants can recall in the correct order is taken as a measure of their STM span (Parkin, 1987). Miller (1956) suggested that STM is constant when measured in chunks or subjectively meaningful units. Baddeley (1990) suggests the existence of an alternative time-based model whereby the main memory capacity, the central executive, dominates over subsidiary systems such as a time based phonological store which is capable of holding information for up to two seconds. Storage is mediated by sub-vocal rehearsal. Memory span is dependent on the rate of rehearsal (Hulme & Tordoff, 1989; Posner, 1964) and recall is a function of how long it takes to articulate the events in memory (Baddely & Hitch, 1974; Baddeley, 1990).

Psychology of Musical Structure

The perception of musical structures such as metre, rhythm, pitch and phrase, have been studied by a number of researchers. The perceptual reality of musical metre has been demonstrated by a number of researchers using tasks such as beat tapping. The main finding is that the highest number of taps occur on position which correspond to the first beat of the bar (Vos, Dijk & Schomaker, 1994). Others have demonstrated that metre is an important memorable feature of music using recognition tasks (Palmer & Krumhansl, 1990). Listeners were better at recognising events which occurred in metrically strong positions of the rhythmical sequence. Jones, Boltz & Kidd (1982) demonstrated that memory for pitches which occurred at positions of metrical importance was better than memory for pitches in weak metrical positions. Smith & Cuddy (1987) found that pitch detection performance was determined primarily by the metre of the musical sequence. In addition to this, it has been demonstrated, using a tapping task, that metrical rhythms are easier to reproduce than non-metrical sequences (Povel, 1981). This type of tapping task has also been used to examine rhythm perception of more complex rhythmical sequences such as polyrhythms which contain a more complex and multi-layered note - note ratio relationship than the sequences typically employed in previous rhythm studies (Handel & Oshinsky, 1981; Handel, 1984). Analyses of piano performances have also revealed that metrical structure is emphasised by accenting the first beat of each bar in comparison to the remaining beats in the bar (Drake & Palmer, 1993).

The importance of rhythm for memory of musical sequences is undeniable. The effect of rhythmic grouping on facilitating STM for a variety of auditory sequences has been demonstrated. The common finding is that recall improves when the material to be remembered is organised or grouped in a rhythmical way (Johnson, 1970; Huttenlocher & Burke, 1969; Bower & Winzenz, 1969; Frankish, 1989; Waterworth, 1983). Rhythm in music acts to highlight points of melodic importance by accentuating important elements (Jones, 1976) and provides a temporal framework to communicate the melody (Handel, 1984). The importance of the temporal organisation in musical sequences, on for example pitch recall, has been demonstrated in number of studies (Deutsch, 1980; Dowling, 1973; Boltz & Jones, 1986). The general finding is that memory is better for sequences in which temporal and tonal patterns coincide in a meaningful way than sequences which contain incompatible temporal and tonal patterns. The way in which pitch and rhythm features are perceptually integrated has also been examined in simple melodies and basic musical stimuli using a detection of change task (Thompson, 1993; 1994). Listeners are sensitive to the way in which pitch and temporal elements are combined in musical sequences, although this effect seems to be dependent on ensuring that attention is focused on the musical sequence.

Pitch structure is another musical feature whose perception has been studied. One of the most basic techniques which has been used to examine musical pitch perception is the probe tone method developed by Krumhansl & Shepard, (1979). This task requires listeners to perform a goodness of completion rating of a sequence of notes. A final note is played after an incomplete ascending or descending note sequence and listeners rate how well it completes the sequence. The basic findings are that listeners rate the tonic of the sequence as the best note to complete the sequences, then other scale members, and non-scale notes are rated as the poorest notes to complete the sequence (Krumhansl & Shepard, 1979; Krumhansl & Keil, 1982; Speer & Meeks, 1985). The psychological importance of the musical scale can be explained by the way in which scales are constructed. Scales commonly contain a limited number of distinct pitch steps, normally 5 or 7 per octave (Krumhansl, 1990; Dowling, 1978). The number of pitch steps correspond to the limited amount of information which listeners can process at a given time (Dowling, 1978; Miller, 1956; Krumhansl, 1990). Neuropsychological research also emphasises the importance of tonal knowledge as an aid to organising melodies into manageable and memorable units (Peretz, 1993). Experimental evidence showed that brain damaged patients, whose impairment disrupted access to tonal knowledge, find it difficult to remember melodies which are longer than six notes. The results imply that, as opposed to non-brain damaged listeners who encode pitch along a tonal dimension to reduce the processing load, brain damaged patients cannot encode pitch without tonal knowledge.

Pitch contour has also been found to be of importance for STM of unfamiliar atonal melodies (Dowling & Fujitani, 1971) and short tonal melodies (Dowling, 1978). Listeners performed a same-different task with a number of transposed melodies which were either exact transpositions, had the same contour but different sized pitch intervals or contained an altered pitch contour. They found that if the contour of the original melody was preserved melody discrimination was poor. Listeners seemed to be representing melodies by contour and focusing less on the exact pitch interval pattern in the melodies. Exact pitch interval sizes are more important in memory for well known melodies (Attneave & Olson, 1971; Dowling & Bartlett, 1981) whereas contour is more important for STM of unfamiliar melodies. The importance of contour as a memorable structural feature in melodies has also been demonstrated in infants (Trehub & Trainor, 1993). These effects have been explained on the grounds that contour information is immediately available to listeners, regardless of whether or not a key has been established or if the melody is familiar to the listener. Interval information on the other hand is inaccurate at the start of a melody before a key is established. This improves as a key is established and is not forgotten as easily as contour information is (Edworthy, 1985). Contour information is most useful in melodies with a weak or confusing tonal context (Dowling, 1978).

The psychological reality of musical phrases has been demonstrated by Sloboda & Gregory (1980) and Stoffer (1985), who have shown that listeners tend to process information within a phrase as a discrete unit of information. It has also been shown that adults and infants are better at detecting pauses within a phrase than between two phrases (Thorpe, Trehub, Morrongiello & Bull, 1988; Thorpe & Trehub, 1989). In addition to this, information within a phrase is remembered more successfully than information across two phrases (Dowling, 1973; Tan, Aiello & Bever, 1981). Memory for phrase structure has been studied in musicians using a melodic recall task. The main finding was that memory of phrase structure in simple melodies is fairly accurate with a tendency for structural simplification of the original phrase structure (Sloboda & Parker, 1985). It has been shown that infants as young as 4 - 6 months are sensitive to phrase structure in music (Jusczyk & Krumhansl, 1993). Infants prefer musical excerpts which preserve the temporal information within each phrase compared with the same excerpts which contain altered temporal information within a phrase (Krumhansl & Jusczyk, 1990). Structural features which were identified as important to musical phrase perception include long duration of last note in a phrase and a drop in the pitch towards the last note. Performance studies have explored the general slowing down in the tempo to signal a musical phrase boundary (Shaffer & Todd, 1987; Repp, 1992).

Research Limitations

Despite the importance of rhythm in perception and cognition of music, rhythm has not been given the attention it deserves in empirical studies. Rhythm was one of the earliest musical structures to be examined by experimental psychologists such as Bolton (1894) and Woodrow (1909) whose interests lay in the perceived grouping of rhythmic sequences. By the 1960's rhythm research was almost non-existent resulting in the term rhythm being removed from the psychological abstract subject index (Handel, 1984). The relative lack of empirical attention directed towards rhythm was also pointed out almost twenty years ago in a chapter lightly titled *Rhythm: tonality's poor relation* in a general psychology of music text (Davies, 1978). The lack of consideration of rhythm as one of the most important musical structures may well be related to the fact that music in the Western classical tradition is often rhythmically simple. The relative imbalance of empirical studies using pitch features and neglecting rhythmic structure is still prevalent today. Most psychology of music texts still devote the main space to studies involving pitch factors. A number of important journal articles have examined memory for the pitches of musical sequences using equitemporal sequences as the experimental stimuli. The so-called rhythm structure is imposed using pitch accent structures (Deutsch & Feroe, 1981) or is induced using temporal gaps which are positioned in an otherwise equitemporal sequence (Dowling, 1973). Rhythmic structure is varied by moving the position of the temporal gap within the sequence.

Another problem is that wide use is made of rigorously controlled experiments which rely exclusively on the most basic short auditory stimuli rather than using musically structured stimuli of greater complexity (Sloboda, 1985; Aiello, 1994). Rhythmic sequences without pitch are employed to study rhythm perception both empirically (Handel & Oshinksy, 1981; Povel, 1981; Povel & Essens, 1985) and theoretically (Longuet-Higgins & Lee, 1982; Lee, 1991; Desain & Honing, 1997; Povel, 1984; Rosenthal, 1989). There has been a tendency to rely upon using a limited number of computer generated sequences to examine perceptual responses to auditory stimuli (Deutsch, 1980; Dowling, 1973). The predominant use of such artificial experimental sequences seriously limits the scope for generalisation of results to a broader musical context.

The role of the non-musical participant

One of the most serious issues in contemporary music perception research is the issue of non-musicians as participants in experiments. Non-musicians are given a raw deal in the majority of music perception research. There is an imbalance in empirical research, with most of the attention directed towards the perceptual responses of the expert participant to music (Smith, 1997). The common trend of much music perception research is to use non-musicians as the predictably incompetent group in empirical studies. This does seem rather like comparing the performance of a group of Olympic standard sprinters with sprinters chosen randomly from the public. Responses of musicians are compared with non-musicians on tasks such as interval discrimination (Burns & Ward, 1978), melody recognition (Dowling, 1991) and completion of pitch sequences (Krumhansl & Shepard, 1979). The results of these task are not surprising: musicians perform significantly better than non-musicians. Do the results tell us something of interest? Is it the case, for example, that musicians are sensitive to the diatonic scale whilst non-musicians are not? (Krumhansl, 1990).

Another interpretation is that non-musicians and musicians interpret a given set of experimental instructions differently (Cook, 1994; Crowder, 1993). Do non-musicians understand the task instructions? Are the statistical differences between the two participant groups due to the two groups engaging in different tasks in response to a given stimulus? There are clear problems with comparing the results from musicians and non-musicians. Another problem is that non-musicians cannot explicitly perform the wide range of musical tasks which musicians can perform. Using musicians as participants gives the experimenter a wider range of tasks to choose from. There are tasks which only those with musical expertise can perform, such as recall of melodies whereby participants recall the melody by writing the music they have just listened to (Sloboda & Parker, 1985; Deutsch, 1980; Boltz & Jones, 1986). Respondents obviously need some form of musical expertise in order to perform such a recall task.

The above discussion has identified a number of features surrounding the music perception field. There are problems of an inability to generalise results of experiments which use impoverished basic experimental stimuli to the actual process of music listening. The imbalance of empirical research on musicians compared with non-musicians demonstrates the need to develop more experimental tasks which are clearly understood by non-musicians and can be used to examine the perceptual responses of non-musical participants. The failure to regard rhythm as an important musical structure in much research must be addressed. There is a clear need for those studying music perception to do so using real musical stimuli as experimental material and to consider results in relation to structural features of the musical stimuli.

The present research

The questions

This research is approached from an experimental psychology angle. Psychological responses to a number of structural features in music are investigated. The experiments examine the perceptual responses of non-musicians to real musical stimuli which contain complex rhythm patterns.

The experimental task

The present research continues the empirical study of music listening. The experimental stimuli are presented to listeners who then perform a given task. Results are analysed and responses are considered in relation to the listeners' perceptual processing capacities and to the pitch and rhythm structures in the stimuli. The types of tasks which are used to investigate the listeners' perception of structure in musical sequences in the present thesis are introduced below.

One chunk or two?

The first pair of experiments in this thesis address the issue of perceptual segmentation in Gaelic melodies. Of interest are the positions at which listeners segment the melodies into smaller units and how these units effect the processing of musical information. The two tasks which examine these points are a segmentation and click localisation task. The segmentation task requires the listener to perform a button-press response as the melody plays to indicate their point of perceived segmentation. The click localisation task is a more subtle way of examining how information in the meaningful chunks is processed. This involves the listener hearing a click in the melody as it plays and then pressing a button to indicate where they heard the click occur during a second playing of the melody. The direction and extent of click migration can highlight perceptual units in the melodies. These two tasks are discussed in greater detail in Chapter 2.

How much can you remember?

Of related interest, is an examination of listeners memory for these unfamiliar Gaelic melodies. Do listeners remember meaningful chunks of information? Do they remember an amount which is equivalent to a perceptual segment in a melody? How do pitch and rhythm features affect memory? Chapter 3 describes a rhythm recall task which was developed to examine these points. The interest lies in examining how much rhythm information listeners could extract after one hearing of an unfamiliar melody.

Can you feel the beat?

Chapter 4 reports the results of a beat-tapping task which was used to examine beat perception in relation to structural features of Gaelic melodies. How quick are listeners to perceive the beat? Do they maintain the beat throughout the course of their listening? These aspects were studied to provide information about beat information in the melodies, specifically how it is communicated and used by listeners as they attend to the melody.

Non-traditional technique

Listeners are required to make some kind of response in experimental tasks in order that music perception can be examined. The type of task which is reported in the chapter 5 differs from the other four experimental tasks used in the present thesis. The task is a two-alternative forced choice task which extracts a forced response from listeners who must select one adjective out of each pair from a set of adjective pairs. Listeners hear a musical fragment and then perform a forced choice response on a list of paired adjectives. This response is not explicitly related to any particular structural feature of the music. There are no detailed instructions to direct the listeners' response and listeners are made aware that there are no right and wrong responses. The task was initially developed to examine the perception of meaning and content in large scale musical forms in an unobtrusive and unbiased manner (Watt, 1995). Participants are not aware that this is what the task aims to do. The present research uses the task to investigate responses to musical intervals and Gaelic melodies.

The experimental stimuli

Real music - the Gaelic melody

The decision to use unaccompanied melodies as the main experimental stimuli was mainly due to the importance of pitch and rhythm structures on the resulting perceptual experience of the melody. Pitch and rhythm are the two musical dimensions which are essential to melody (Boyden, 1971). Melodies are highly organised and structured sound sequences containing events which change over time (Pick, 1979). It is the relations between the various pitch and rhythm events which are important to the perceptual experience of the melody. Listeners who are unfamiliar with a given melody can still perceive properties of the structure when listening to the melody for the first time (Dowling, 1978). Melodies are thus an ideal auditory sequence with which to examine the listeners' perceptual and cognitive responses to music. The melodies used in the present experiments are mainly melodies taken from *Bineas Nam Bard*, a selection of Scottish Gaelic folk songs. These melodies are richly expressive in an unaccompanied form as Gerstley (1997) confirmed by studying them in a perceptual-emotional quality experiment. Listeners heard a melody and then performed an emotion perception task by positioning a sliding pointer on a computerised visually displayed *Happy-Sad* scale. Melodies were classified into sad, neutral and happy groups as determined by the consistency and direction of the emotional ratings. Many of the melodies contain a strong emotional content which is communicated in one single line of music. The melodies served three main purposes for the present set experiments:

1. The majority of listeners were unfamiliar with these melodies which is essential for experiments aimed at examining processing units and immediate short term memory of the rhythmic structure of the melodies.
2. They were classifiable into three main emotional or mood related categories, sad, neutral and happy which was ideal for experiments which examine information units and consider the content and meaning of melodies.
3. The Gaelic melodies used in the present research contain a well-structured set of complex musical rhythms. These rectify the tendency to use rhythmically dead stimuli such as equitemporal tone sequences which have been predominantly used in previous research.

The need for controls

Experimental studies in this thesis make use of Gaelic folk tunes in the main plus isolated two note intervals and computer created sequences which contain rhythm and pitch patterns which are generated by a set of probabilistic rules. All experimental stimuli were produced via computer and were timed accurately. Thus melodies were produced without dynamic accents and without expressive timing and dynamics. Computer timed melodies are a suitable material with which to examine perceptual responses. It is only possible to explain perceptual responses in terms of basic structural features in the music such as pitch and rhythm, without the additional confounding factors of expressive timing deviations (Repp, 1992; Shaffer & Todd, 1987) and accenting of notes (Drake & Palmer, 1993) so common to natural musical performances. This kind of control was especially important because the present research methods employed a number of novel techniques with which controlled stimuli is a necessary starting point.

The author acknowledges that there are a number of important features in expressive performances of music. Indeed recent empirical research has demonstrated the effect of the performers expressive intentions and individual differences between performers of Swedish folk tunes on listeners' perceptual responses to those tunes (Gabrielsson &

Juslin, 1996). The problem with using expressive performances of the music is that musical performances can vary depending on who is playing or singing the melody, on what instrument they are using or which particular voice range they sing in and on their expressive intentions. The Gaelic melodies used in the present study are richly expressive and often contain strong emotional content. Rather than include potentially confounding expressive factors in the present experimental stimuli, experiments make use of the melodies as they are written.

Summary

Each of the tasks used in the thesis examine a different, although interrelated, aspect of the musical structure of melodies. The segmentation and click localisation work in chapter 2 focuses primarily on perception of phrasal units in the melodies. The recall experiments of chapter 3 deal specifically with memory for rhythmic structure, although pitch is an equally important structural feature in this study. The beat perception research in chapter 4 focuses specifically on the musical beat of the melodies. The issue of content and meaning in the melodies is examined in chapter 5. The research examines the perception of melodies at the level of the phrase, of rhythmic groups, of the beat and finally addresses the perception of content and meaning in the course of listening to the melodies.

Chapter 2 Segmentation and Click Migration in Gaelic melodies

Experiment I Segmentation of Gaelic melodies

The starting point of the present experiment was to identify units in the Gaelic melodies which listeners perceive as discrete segments of information and to consider what structural factors in the stimuli may be shaping these perceptual groupings. It was important to establish the occurrence of perceptual or phrasal groups in the melodies for all subsequent experiments including the click migration experiment which is reported in Experiment II of the present chapter. As discussed in chapter 1, the experiments in the present thesis were developed to examine different aspects of the listeners perceptual and cognitive responses to music. Although each chapter focuses on a different aspect of musical structure, the experiments relate directly to the information units into which listeners segment the melody. Hence the importance of the segmentation experiment. The first section of the present chapter begins with a description of previous research on perceptual grouping in auditory sequences. Empirical work which examines segmentation of music is considered before the present segmentation work is described.

Perceptual groups

Listeners perceptually group incoming auditory information. The tendency to group auditory information is so strong that listeners perceive accents in auditory sequences when there are in fact no accents present (Bolton, 1894; Fraisse, 1982; MacDougall, 1903; Povel & Okkerman, 1981; Temperley, 1963). The importance of grouping is reflected by the large number of studies which have examined the effects of grouping in experimental sequences on recognition, discrimination and immediate memory of listeners (Martin, 1972; Handel & Oshinsky, 1981). The fundamental idea behind such work is that the objective groups which correspond best to the listeners perceptual groups will facilitate performance on a given task. An examination of the grouping processes which listeners engage in should provide insights into the way in which auditory information is processed.

Given that listeners perceptually group incoming information, what are the factors which determine where subjective grouping will occur? The importance of pauses in determining segmentation of sequences has been widely documented. It has been demonstrated that pauses in sequences of lights and digits provide perceptual groupings (Restle, 1972; Bower & Winzenz, 1969). Studies have demonstrated the importance of temporal pauses in segmenting, and facilitating our memory for, lists and spoken sentences (Frankish, 1989), telephone numbers (Waterworth, 1983), as well as structured tonal sequences (Deutsch, 1980; Dowling, 1973).

A great deal of research has examined the way in which perceptual groups effect the processing of information in auditory sequences. Tasks such as detection of pauses within and between groups in a sequence are commonly used. Thorpe, Trehub, Morrongiello & Bull (1988) examined discrimination of temporal variations in auditory sequences by infants. Infants heard tone sequences which contained pauses which were positioned within or between the groups of a sequence. The results demonstrate that infants were poorer at detecting the pauses which were positioned between the groups compared with those positioned within the groups in the auditory sequences. Thorpe & Trehub (1989) report similar results. Pauses which were placed within groups of tones were more successfully detected than pauses which were placed between groups of tones. A similar pair of studies were carried out by Jusczyk & Krumhansl (1993) and Krumhansl & Jusczyk (1990). Using a habituation method, they showed that infants preferred musical excerpts which contained pauses which were inserted at phrase boundaries compared with excerpts which contained pauses which were inserted in the middle of the phrase. Pitch factors have also been considered in perceptual grouping experiments. The main finding is that pauses which were positioned between tones of similar frequency were detected more correctly than pauses positioned between tones of dissimilar frequency (Fitzgibbons, Pollatsek & Thomas, 1974). These results demonstrate that listeners perceptually differentiate between groups of elements in an auditory sequence with the use of temporal and pitch cues. The differentiation between groups of tones occurs to the extent that pauses between groups are perceived to occur naturally and therefore go unnoticed whereas pauses within a group are detected.

So why do listeners group? The answer to this question may be related to limitations in listeners' perceptual processing capacities. The amount of incoming information which listeners can deal with at a given point in time is limited. Grouping allows listeners to subjectively chunk a given sequence into smaller units and thus differentiate subgroups in the sequence from others. Perceptually grouping an auditory sequence may facilitate efficient processing and enhance memory for the auditory information. It has been demonstrated that perceptual groups can be experimentally manipulated by simple alterations to stimuli such as temporal pauses within a sequence (Frankish, 1989; Dowling, 1973). In addition to this, the detection or tolerance for such pauses is dependent on whether the pauses occur within or between groups (Thorpe & Trehub, 1989; Jusczyk & Krumhansl, 1993). This reflects the tendency of listeners to hear groups within a sequence as discrete and distinct from other groups in a sequence. It is a natural progression to consider how listeners segment musical sequences into groups and to consider what the function of these groups may be.

Musical phrases

The importance of phrases to music perception is undeniable. Segmenting a musical piece into smaller units such as motives, phrases and sections, arranged hierarchically, is of vital importance for understanding music (Jackendoff, 1991). Aiello (1994) suggests that listeners segment language and music in locations which reflect grammatical structure and proposes that a phrase is a psychological unit in both language and music. Swain (1986) supports the idea of a musical 'clause' or basic processing unit and natural unit of meaning. Sloboda (1977) provides empirical evidence that the phrase is a distinct unit of meaning which directly effects the memory span of sight-readers when performing. The perceptual experience of music is dependent on the ability of listeners to segment music into groups (Bigand, 1993). In order to understand music perception, there is a need to know more about the psychological principles involved in the segmenting of sound sequences into orderly meaningful sub-units such as phrases (West, Howell & Cross 1991).

The lack of theoretical and empirical research carried out on the processing of phrases has been pointed out by many such as West et al (1991) who suggest that the lowest (e.g. notes) and highest (e.g. movements) levels of identification of musical subunits have been examined and argue that it is at the intermediate level of the phrase that we remain the most ignorant. Sloboda and Gregory (1980), argue that we do not have an adequate explanation of how listeners segment a melody into phrases. In addition to this, Todd (1994) argues that computational models of musical rhythm have largely neglected phrasal grouping.

Many experiments have investigated how listeners process information within and across musical phrases with the general result that within phrase information is more successfully remembered and utilised in test situations. Tan, Aiello & Bever (1981) used a two-note probe recognition task to examine recognition of tones which occurred within or between a phrase in a melody. Listeners heard the melody which was then followed by a two-note probe. The probe consisted of two notes which: ended the first phrase; crossed the phrase boundary or began the second phrase. The task was to indicate if the two-note probe had occurred in the melody. The results showed that the probe which cut across two phrases was the most difficult to recognise. Dowling (1973) also demonstrated the importance of phrasal information with a recognition task which was designed to study memory for five tone sequences. Listeners heard a twenty tone sequence which was rhythmically grouped by four groups of five tones. Memory of test sequences taken from a complete group of the twenty tone sequence (within group) was better than memory for test sequences which were taken from the end of a group and the beginning of another in the twenty tone sequence (crossed between two groups). The importance of temporal factors in determining the formation of perceptual groups and memory of simple tonal sequences has also been demonstrated by Deutsch (1980). She reports that sequences of tones are recalled more successfully when temporally segmented into chunks which are compatible with the tonal structure compared with chunks which are incompatible with the prevailing tonal structure.

This experimental evidence demonstrates that it is easier to process and remember information which occurs within phrases compared with information which occurs across two phrases. Recognition of notes which occur within a phrase is better than recognition of notes which occur across two phrases (Tan et al, 1981; Dowling, 1973). The results of the above experiments suggest that there is some form of discrete perceptual unit in music processing which is maintained in memory and used in such memory tasks. The amount of incoming information which listeners can handle is limited. We can suppose that listeners have some kind of maximum level of information processing capacity which is mediated by a variety of factors such as chunking or temporal segmentation of the sequence into smaller more meaningful units. We may also presume that such information is preserved within these groups as chunks and can actively disrupt or enhance performance on memory tasks depending on whether the test stimuli have been grouped in accordance with chunks which listeners perceive.

There has also been a resurgence of interest in Gestalt principles (Michon, 1985), a set of basic laws which were proposed to explain perceptual groupings of visual stimuli (Wertheimer, 1927, as cited in Gleitman, 1986). Grouping principles such as proximity (figures close to one another are likely to be perceptually grouped together), closure (the tendency to perceptually complete figures containing gaps) and continuity (preference for smooth continuous contours) were used to describe preferences for well-formed structure. Although applied to visual patterns, these principles are relevant to the auditory domain (Trehub & Trainor, 1993) and can be used effectively to describe perception of local patterns in music (Dowling & Harwood, 1986). Gestalt laws have been tested in relation to grouping and segmentation of music in an attempt to examine how listeners parse music into meaningful perceptual units (Deutsch, 1982b; Sloboda, 1985). It has been suggested that considering gestalt perception in relation to music allows us to understand how listeners categorise and chunk sequences into sub-units such as melodic phrases (Terhardt, 1987).

It has been noted that much of the experimental stimuli used in these phrase perception tasks were short artificial sequences. The sequences used by Deutsch (1980), Dowling (1973) and Tan et al (1981) lacked any realistic rhythmical component which real music would have contained. The pitches of experimental sequences used by Deutsch (1980) were generated according to tightly controlled rules which were tested by an extremely limited number of sequences. Dowling (1973) utilised sequences which were generated by a probabilistic set of ups and downs, favoured minor seconds half the time and major seconds and minor thirds for the rest of the time and contained no other pitch interval.

Musical phrase perception

How do listeners actually perceive a phrase in music? Krumhansl (1996) argues that music consists of units, with well defined beginnings and endings, in which new elements are introduced, developed and defined by a variety of acoustic cues such as pauses, pitch contour and rhythmic patterns. Attempts have been made to identify the underlying factors which contribute to the listener's sensitivity to phrase structure in large scale musical forms. The results show that the segmentation of music depends on a multitude of factors such as melodic contour changes, lengthening of duration's, pitch range contrasts, dynamics, timbre, tempo of performance and meter (Deliege, 1987; Todd, 1985; Deutsch, 1980; Clarke, 1982; Clarke & Krumhansl, 1990; Parncutt, 1994; Benjamin, 1984). As mentioned in Chapter 1, experiments which were carried out with infants identify a drop in pitch height and an increase in tone duration prior to a pause in the sequence as the most important indicators of phrase structure (Jusczyk & Krumhansl, 1993). Others have demonstrated that expressive performers can aid phrase perception. Drake & Palmer (1993) argue that metric and melodic accent structures and rhythmic grouping influence the listeners segmentation of a musical sequence. They report that performers enhance such accent structures to aid the listeners segmentation process.

Segmentation in music

Segmentation of music has a predominant role in Lerdahl & Jackendoff's (1983) theory of the final stages of musical comprehension in Western tonal music. They propose segmentation rules to account for the ways in which the musical surface is split into a series of time-spans. Time-spans are described as rhythmic units which are understood with respect to pitch structure. The nature of the time-span in which the music is segmented depends upon the level in the music under discussion. At the local level, below the level of the phrase, time-spans are determined by metrical structure and an interaction between metrical structure and grouping. In the intermediate level, time-spans refer to uncadenced groups and in the final phrase level, time-spans refer to cadenced groups. Theoretically, the phrase is important to the final stages of musical understanding. However there is little empirical evidence to support the proposed grouping structures.

Empirical research has tested Lerdahl & Jackendoff's (1983) "Preference rules" which attempt to account for the groupings which listeners perceive in a given musical sequence. Deliege (1987) examined the listener's perceptual segmentation of a variety of different musical excerpts such as Baroque, Classical, Romantic and nine note sequences which were specifically constructed to contain conflicting grouping principles. The main task was to segment the music by marking the perceived phrasing on a simplified musical score which consisted of a series of equally spaced dots. Results showed that points of perceived segmentation supported predictions which were made were by Lerdahl & Jackendoff's (1983) grouping rules. However, criticisms have been made about such unnatural forms of experimental task. It is possible that experiments which demand such an unusual listening response may alter the perceptual experience of the listener (Cooke, 1994). It is likely that the segmentation responses in Deliege's (1987) study do not reflect the normal listening process.

Similar experiments have been performed to examine segmentation of large scale musical forms, using tasks which interfere minimally with the normal listening process. Clarke & Krumhansl (1990) studied segmentation of atonal and classical music using a task which required the listener to press a foot pedal after hearing a boundary. This avoids any visual distraction that following a notated score, simplified or otherwise, may have caused. A wide variety of structural features in the music were found to affect perception of boundary occurrence, such as silences, long pauses, pitch content and melodic contour and changes in rhythm, dynamics and timbre. They also report that their findings are in accordance with predictions of Lerdahl & Jackendoff's (1983) grouping preference rules. Krumhansl (1996) used a similar task to examine segmentation of Classical music. Listeners were required to press a computer mouse button when they heard each major section occur. Performed tempo was the strongest indicator of perceived phrase boundary with slowing down occurring at major phrase boundaries. In contrast to Clarke & Krumhansl (1990) and Deliege (1987), Krumhansl's (1996) segmentation results did not correspond well to predictions made by Lerdahl and Jackendoff (1983).

The present research

There has been a great deal of empirical work carried out to examine grouping principles, segmentation and processing of information within and across phrases. In addition to this, there are some well defined theoretical explanations which attempt to account for perceptual groupings of musical sequences. However, research based around the musical phrase is severely lacking in a great many respects. The majority of empirical research has focused on either large scale musical pieces or short, artificial sequences. There is no neat explanation or consensus of what constitutes a psychological phrase when listening to real musical forms, nor are there many attempts to quantify the nature of perceptual groups in relation to the limited cognitive capacities of listeners. It is necessary to broaden our understanding of what constitutes a musical phrase or information unit, find out more about how listeners process phrasal information and examine the listeners cognitive capacities for dealing with incoming auditory information in a quantifiable amount, such as time in seconds.

Given the problems of using short artificial sequences (inability to generalise the results) and large scale musical pieces (such a multitude of factors can influence phrase perception), it was deemed necessary to reach some middle ground with the experimental stimuli in the present segmentation experiment. Unaccompanied real melodies, without expressive timing deviations or dynamic variations were utilised for these reasons. The segmentation task was used to:

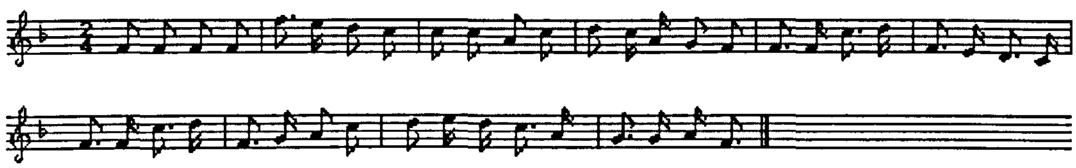
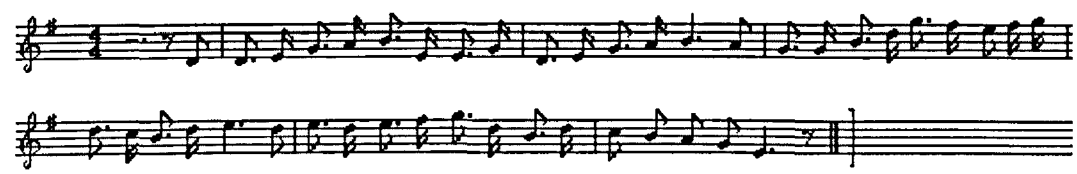
1. Identify where listeners perceptually segment the melodies and thus highlight the occurrence of these subjective units of information for subsequent experiments and
2. Identify structural features in the melodies which may be responsible for eliciting the segmentation response.

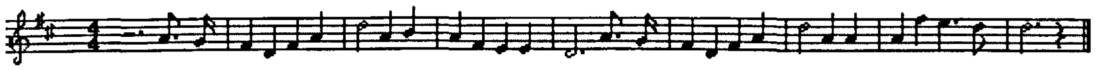
Method

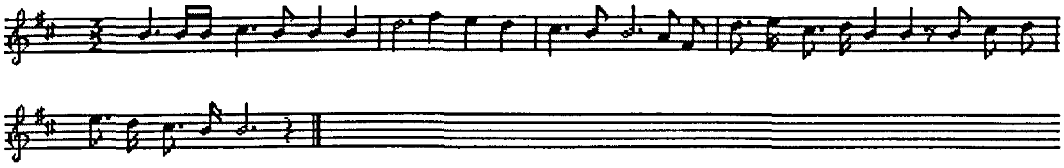
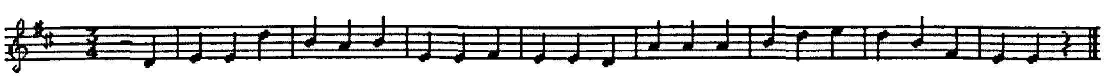
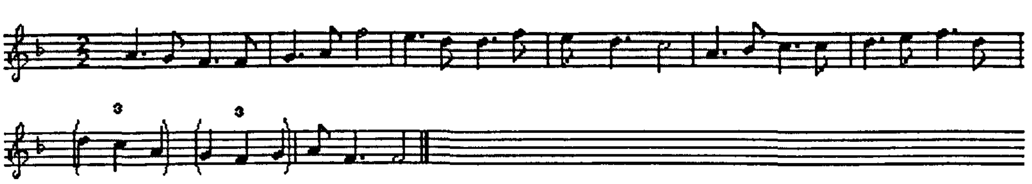
Subjects The subjects were 119 undergraduates, 68 female, 51 male, aged between 18 and 36 years. Participation in the experiment was in part fulfilment of a psychology course requirement.

Stimuli/Apparatus The stimuli consisted of excerpts from six Scottish Gaelic melodies, h1, h2, n1, s1, s2 and s3, shown in Figure 1, below. The melodies were selected from a large set that had been previously rated on a perceptual-emotional quality scale (Gerstley, 1997). Two melodies from the consistently rated *happy* category (h1 and h2), three from the *sad* category (s1, s2 and s3) and one from the *neutral* category (n1) were used in the present experiment.

Figure 1. Melodies used in the Segmentation Experiment.

Melody H1

Melody H2

Melody N1



Melody S1

Melody S3

Melody S2

Melody S2 unstructured

The melodies were performed on a computer without expressive performance deviations. This provided the advantage of enabling the experimenter to explain perceptual responses in relation to particular structural properties of the stimuli. The rhythm and pitch structures were complex and varied in most of the melodies. S1, S2 and N1 contained longer note durations than H1 and H2. S3 contained only crotchets, there were no different durations in the piece. Melody S2 unstructured was the first two phrases of melody S2 randomly rearranged to examine if segmentation responses would occur regardless of note order. These were produced on a NeXT using a bagpipe timbre. They were played at a tempo of one crotchet beat per 0.60 second without expressive timing and dynamic variation. Listeners' responses (onset and duration) were detected by an electronic switch and recorded on a Sound Edit package on a Macintosh Centris 650.

Procedure Listeners were asked to segment the melody into meaningful units by pressing an electronic hand-held switch when they thought that a meaningful unit had occurred. They were instructed to listen to the melody excerpt three times without actively responding and then to practise segmenting the excerpt into meaningful units by depressing the switch when they thought a meaningful unit had occurred. They were asked to practise until they were satisfied with their responses. Their responses were then recorded. 24 of the participants performed with melodies H2, N2 & S2, 24 performed with melody S1, 34 with melody H1, 18 with melody S3 and 19 with melody S2 unstructured. The experiment lasted an average of 20 minutes. Listeners completed a short questionnaire which explored musical background and familiarity with the melodies used in the experiment.

Results

Figures 1 - 7, below, show the segmentation responses which are grouped into quaver units of time (0.30 second blocks of time). It was assumed that perceptual boundaries, like phrase boundaries, would generally occur with the end of a note. The time of occurrence of a note-end was subtracted from the time of button-press to give the position of the perceived point of segmentation in relation to the end of notes with the longest duration in the melody. Thus the distance (how far the segmentation response is from note offset) and direction (earlier, during or after note) can be calculated for each melody.

As shown in Figures 2, 3, 4, 5 & 6, segmentation responses are fairly consistent for melodies H2, N1, S1, S2 and S2 unstructured with the majority of subjects using note duration as the major indicator of phrase boundary occurrence. The majority of responses cluster around the longest notes in each melody. Responses for melodies H1 and S3 are spread throughout the melody rather than clustered into groups. Responses to each melody are discussed below.

Melody h1 Descriptive results

Responses are shown in Figure 1, below. These reveal two distinct response clusters which are grouped around the two quavers which occur after measures 2 and 4 in the melody. 61.80 % of listeners respond within 0.60 seconds of the quaver note ending measure 2 and 85.29 % respond within 0.80 seconds of the quaver ending measure 4. The responses for the rest of the melody are more variable and do not cluster into one particular group.

Duration of segments

The duration of perceived segments 1 and 2 is 2.4 seconds each. In other words, the start of the melody to perceived boundary 1 is 2.4 seconds and from boundary 1 - boundary 2 is 2.4 seconds in time.

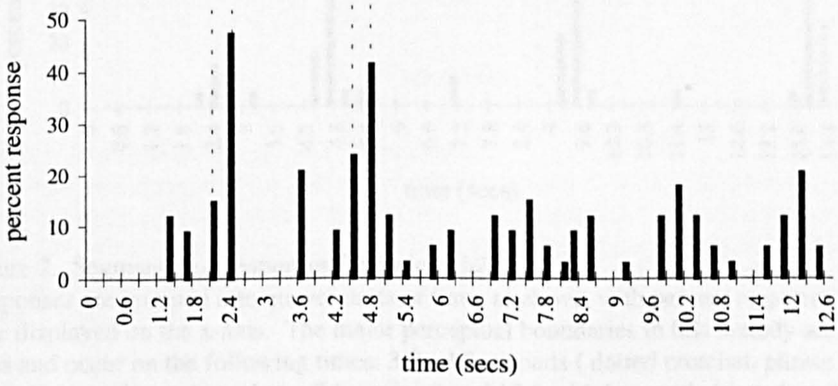


Figure 1. Segmentation responses for melody H1.
Responses are grouped into quaver units of time (0.30 seconds on the x-axis) in relation to note offsets. The y-axis represents the percentage of subjects responding in a 0.30 second (quaver) period of time. The major perceptual boundaries in this melody are marked by the dotted line and occur on the following times: 2.1 - 2.4 seconds (boundary 1, quaver) and 4.5 - 4.8 seconds (boundary 2, quaver).

Melody h2 Descriptive results

Responses are shown in Figure 2, below. These reveal three distinct response clusters which are grouped around the three longest notes (dotted crotchets) in the melody.

Phrase 1

87.5 % of subjects responded as the boundary note sounded (on the note onset) and 8.3 % responded on the note following the long note, within 0.60 seconds of the boundary note offset.

Phrase 2

This phrase elicited similar responses, with 95.7 % of subjects responding as the dotted crotchet sounded. The remaining 4.3 % responded within 0.06 seconds of the note offset.

Phrase 3

The boundary of phrase 3 elicited segmentation responses in 75 % of participants, all of whom responded as the note sounded.

Durations of segments

The duration of the three identified segments was 4.8 seconds. In other words, the start of the melody up to boundary 1 was 4.8 seconds in duration. This was also the duration of boundary 1 - boundary 2 and boundary 2 - boundary 3.

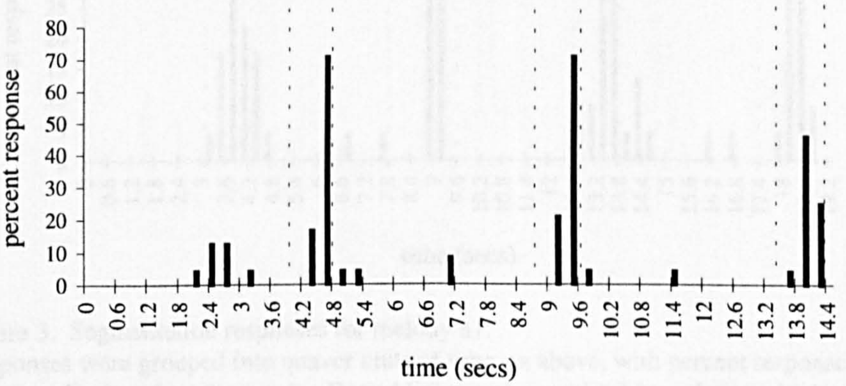


Figure 2. Segmentation responses for melody h2.
Responses are grouped into quaver units of time, as above, with percent response shown on the y-axis and time displayed on the x-axis. The major perceptual boundaries in this melody are marked by the dotted lines and occur on the following times: 3.9 - 4.8 seconds (dotted crotchet, phrase 1 boundary); 8.7 - 9.6 seconds (dotted crotchet, phrase 2 boundary) and 13.5 - 14.4 seconds (dotted crotchet, phrase 3 boundary).

Melody n1 Descriptive results

The results revealed four main clusters of responses which correspond to the longest held notes in the melody, shown in Figure 3, below.

Phrase 1

This phrase elicited responses from 95 % of participants. 70 % responded as the note sounded and 25 % responded within 0.60 seconds of the boundary offset.

Phrase 2

Segmentation responses were similar for phrase 2 with 82.8 % responding as the boundary note sounded and 13 % responding on the note prior to the boundary onset.

Phrase 3

Again, most listeners responded to this boundary note. 74 % responded as the note sounded and 21.8 % within 0.60 seconds of the note offset.

Phrase 4

The final note elicited responses from 62.5 % of participants, all of whom responded as the note sounded.

Duration of segments

The durations of the four identified segments were as follows: start of melody - boundary 1: 4.2 seconds; boundary 1 - boundary 2: 5.4 seconds; boundary 2 - boundary 3: 4.2 seconds; boundary 3 - boundary 4: 5.4 seconds.

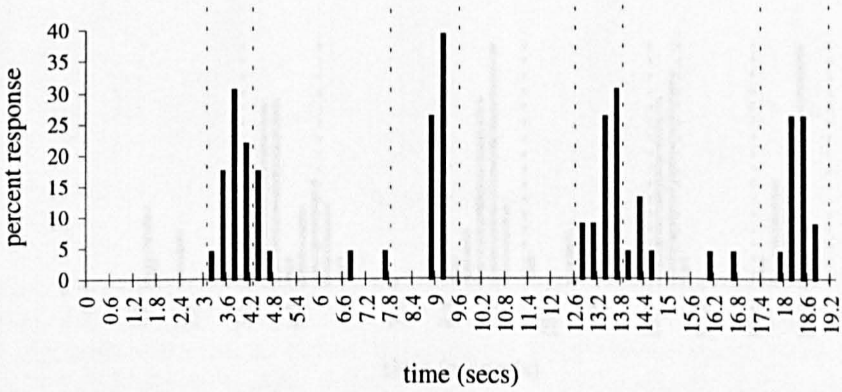


Figure 3. Segmentation responses for melody n1.
Responses were grouped into quaver units of time, as above, with percent response shown on the y-axis and time displayed on the x-axis. Dotted lines mark perceived boundaries which occur on the following times: 3.0 - 4.2 seconds (minim, phrase 1 boundary); 7.8 - 9.6 seconds (dotted minim, phrase 2 boundary); 12.6 - 13.8 seconds (minim, phrase 3 boundary) and 17.4 - 19.2 seconds (dotted minim, phrase 4 boundary).

Melody s1 Descriptive results

The results from melody s1 are shown below in Figure 4. Four main response clusters are revealed. These clusters group around the three longest held notes in the melody (dotted minims) and a quaver rest.

Phrase 1

The first phrase elicited responses from all subjects. 83.3 % responded as the boundary note sounded and 12.5% responded within 0.30 seconds of the boundary note offset. The remaining 4.2 % responded on the note prior to the boundary onset.

Phrase 2

Similar results were found for phrase boundary 2 with 95.8 % of subjects responding as the note sounded, the remaining 4.2 % responding 0.20 seconds after the note offset.

Phrase 3

Phrase boundary 3 elicited responses from 79.2 % of subjects. All responded within 1.2 seconds of the rest onset.

Phrase 4

The final phrase of the melody excerpt elicited responses from 79.2% of subjects, all of whom responded as the note sounded.

Durations of segments

The durations of each of the four identified segments were as follows: start of melody - boundary 1: 5.4 seconds; boundary 1 - boundary 2: 4.8 seconds; boundary 2 - boundary 3: 3.3 seconds; boundary 3 - boundary 4: 3.9 seconds.

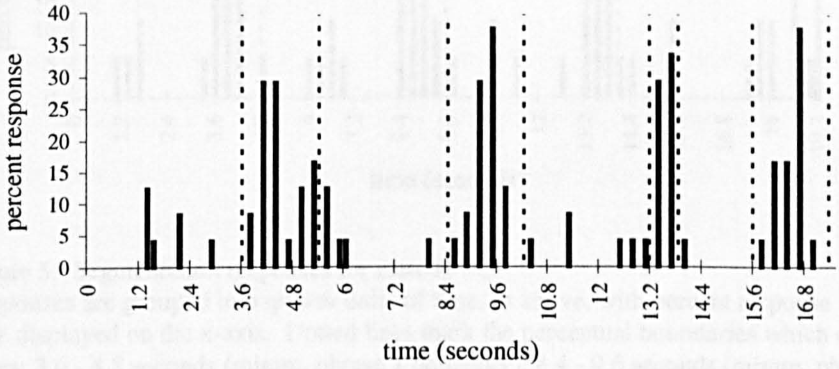


Figure 4. Segmentation responses for melody s1.
Responses are grouped into quaver units of time as above in Figure 1a, with the percentage of responses shown on the y-axis, occurrence in time on the x-axis. Dotted lines mark the major perceptual boundaries which occur on the following times: 3.6 - 5.4 seconds (dotted minim, phrase 1 boundary); 8.4 - 10.2 seconds (dotted minim, phrase 2 boundary); 13.2 - 13.5 seconds (quaver rest, phrase 3 boundary) and 15.6 - 17.4 seconds (dotted minim, phrase 4 boundary).

Melody s2 Descriptive results

The results from melody s2 are shown below in Figure 5. Three main response clusters are revealed. These clusters group around the three longest held notes in the melody (two minims and a dotted crotchet).

Phrase 1

The first phrase boundary elicited responses from 94.44 % of participants. 83.33 % of participants responded as the note sounded and 11.11 % responded within 0.07 seconds of note offset.

Phrase 2

Similar responses were found at the second phrase boundary with 83.33 % of listeners responding as the boundary note sounded and 5.55 % responding within 0.06 seconds of the note offset.

Phrase 3

The dotted minim also elicited a similar degree of responding with 77 % of participants responding as the note sounded and 11.11 % responding within 0.30 seconds of the note offset. 5.55 % responded within 0.60 seconds prior to the note onset.

Duration of segments

Segments 1 and 2 are 4.8 seconds in duration and segment 3 is 4.5 seconds long.

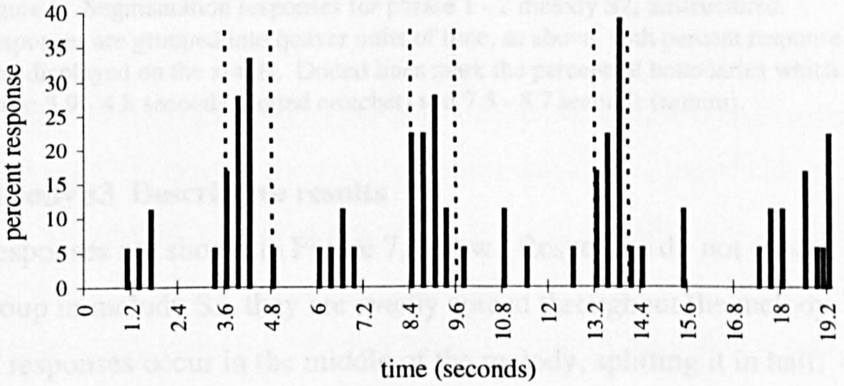


Figure 5. Segmentation responses for melody S2.
Responses are grouped into quaver units of time, as above, with percent response shown on the y-axis and time displayed on the x-axis. Dotted lines mark the perceptual boundaries which occur on the following times: 3.6 - 4.8 seconds (minim, phrase 1 boundary); 8.4 - 9.6 seconds (minim, phrase 2 boundary); 13.2 - 14.1 seconds (dotted crotchet, phrase 3 boundary).

Melody s2 unstructured Descriptive results

The results from melody s2 unstructured are shown in Figure 6, below. Two main response clusters are revealed. These clusters group around the two longest held notes in the melody (dotted crotchet and minim).

Phrase 1

The first dotted crotchet elicited responses from 90 % of participants. 79 % of listeners responded as the note sounded and 11 % of listeners responded within 0.07 seconds of the note offset.

Phrase 2

The minim elicited responses from 68 % of participants, all responding as the note sounded.

Duration of segments

Segment 1 is 4.8 seconds long and segment 2 is 3.9 seconds long.

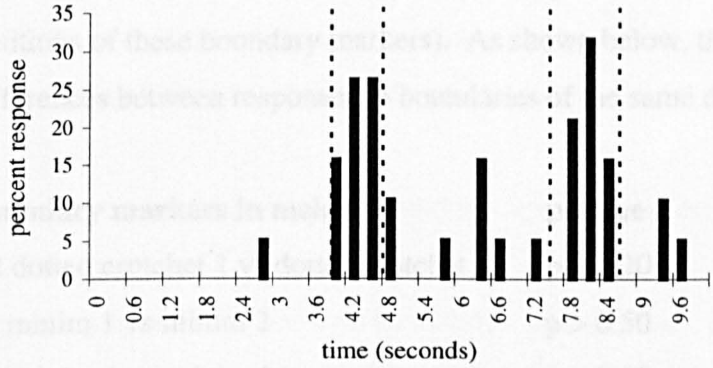


Figure 6. Segmentation responses for phrase 1 - 2 melody S2, unstructured.
Responses are grouped into quaver units of time, as above, with percent response shown on the y-axis and time displayed on the x-axis. Dotted lines mark the perceptual boundaries which occur on the following times: 3.9 - 4.8 seconds (dotted crotchet) and 7.5 - 8.7 seconds (minim).

Melody s3 Descriptive results

Responses are shown in Figure 7, below. Responses do not cluster into any particular group in melody S3, they are evenly spread throughout the melody. The largest group of responses occur in the middle of the melody, splitting it in half. 40 % of participants make a segmentation response at this point in the melody which splits it into two segments of 7.8 and 6.6 seconds in duration.

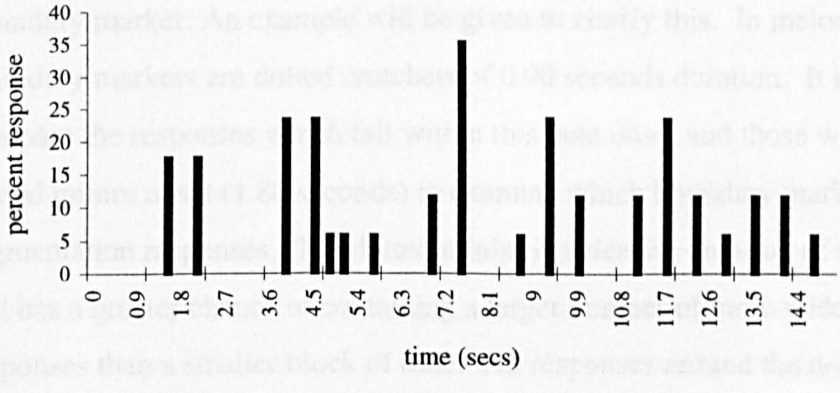


Figure 7. Segmentation responses for melody S3.
Responses are grouped into quaver units of time, as above, with percent response shown on the y-axis and time displayed on the x-axis.

Speed of button-press response

A series of repeated measures ANOVA's were performed on the mean segmentation responses (absolute distance from note offset) for each of the main perceptual boundaries in melodies which contained obvious structural markers (the first two dotted crotchets in H2, the first two minims in S2, the first two minims & dotted minim in N1, and the first two dotted minims in S1. See Figures 2, 5, 3 and 4, above, for the temporal positions of these boundary markers). As shown below, there were no statistical differences between responses to boundaries of the same duration.

Boundary markers in melody	**p value**
H2 dotted crotchet 1 vs dotted crotchet 2	$p > 0.20$
S2 minim 1 vs minim 2	$p > 0.50$
N2 minim 1 vs minim 2	$p > 0.60$
S1 dotted minim 1 vs dotted minim 2	$p > 0.10$

Responses were grouped by duration of boundary marker and melody type for statistical analyses. There were five different boundaries in the three melodies: 1) quaver rest in S1 (0.30 seconds); 2) dotted crotchet in H1 (0.90 seconds); 3) minim in S2 (1.20 seconds); 4) minim in N1 (1.20 seconds); 5) dotted minim in N1 (1.80 seconds) and 6) dotted minim in S1 (1.80 seconds). The area from which segmentation responses were taken was 0.30 seconds prior to the dotted minim onset and 0.30 seconds after note offset (i.e. 2.40 second block of time). In order that responses are comparable in each melody, responses were selected from the 2.40 second block surrounding each boundary marker. An example will be given to clarify this. In melody H2, the two boundary markers are dotted crotchets of 0.90 seconds duration. It is not sensible to compare the responses which fall within this note onset and those which fall on the dotted minim onset (1.80 seconds) to examine which boundary marker elicits faster segmentation responses. The dotted minim is twice the duration of the dotted crotchet and has a greater chance of containing a larger number of more widely spread motor responses than a smaller block of time. All responses around the dotted crotchet *within the same block of time* as responses taken around the dotted minim are included in the analysis. As mentioned above, segmentation responses are selected from 0.30 seconds before and after the dotted minim note onset and offset. To keep the duration of time in which responses are included the same for melody H2, responses which fell 0.75 seconds before the dotted crotchet onset up to 0.75 seconds after the dotted crotchet offset are included in the analysis. This results in a 2.40 second block of time around the dotted crotchet which is comparable to the 2.40 second block of time around the dotted minim. Similarly, responses 0.60 seconds before minim onsets and 0.60 seconds after minim offsets are included, as are responses 1.05 seconds before quaver rest onset and 1.05 seconds after quaver offset. Responses which occur within the 2.40 second block of time around each structural marker are included.

The comparisons between segmentation responses around these five boundaries enabled an analysis of:

1. melody type (S1 dotted minim with N1 dotted minim, and S2 minim with N2 minim, i.e. comparing the same boundary note durations between two different melodies),
2. duration alone (N1 minim compared with N1 dotted minim, S1 quaver rest with S1 dotted minim, i.e. different note durations within the same melody) and
3. general comparisons between melody type and boundary note durations (e.g. S1 dotted minim with H1 dotted crotchet, S1 quaver rest with N1 minim, and so on).

This was carried out to ascertain whether spread of response about the boundary note offset was similar for a variety of boundary note durations (e.g. dotted crotchet versus dotted minim) and different contexts (e.g. melody S2 versus melody N1) or whether the speed of listeners button-press response is quicker with shorter boundary notes and dependent on the particular context (melody). In other words, does length of boundary note determine how far from note offset the listeners responses occur? Is speed of response dependent on the particular melody?

Speed of segmentation results

The mean segmentation responses (distance from note offset) for each of the main boundaries are displayed in Figure 8, below, and the ANOVA results are summarised in Table 1, below. Comparisons of responses around boundary markers of different durations across different melodies revealed that listeners generally respond more quickly to shorter boundary markers. There was a significant effect of note duration for the quaver rest boundary in S1 (the shortest boundary marker) compared with longer boundaries. Listeners are faster to respond to the quaver rest than to boundary markers of a minim or longer. There was a significant effect of boundary note duration for the dotted crotchets in H2 compared with the minim of S2 and the dotted minims of N1 and S1. Again, faster responses are made on the shorter dotted crotchet marker. There was no difference between dotted crotchet boundaries in H2 and the minim boundary of N1. There was a significant difference between responses around the N1 minim boundary and the S1 dotted minim boundary with faster responding on the shorter minim boundary. There was no difference between the S2 minim boundaries and S1 dotted minim boundaries.

Melody type

There was no significant difference between responses to the dotted minims of melody N1 and the dotted minims of melody S1, thus there was no effect of melody type on segmentation responses to the longest dotted minim boundaries. However, there was a significant difference between minim boundaries of S2 and minim boundaries of N1 with faster responding to the minim boundaries of N1. No other pure melody comparisons could be made because no two melodies contained a boundary marker of the same duration.

Note duration

A significant effect of note duration was revealed by the comparison of responses to boundary N1 minim and N1 dotted minim, with faster responses on the shorter boundary compared with the longer boundary. There was also a significant difference between boundary S1 rest and S1 dotted minims with faster responding on the shorter quaver rest. No other melodies contained two different boundary durations.

Table 1 Anova results of segmentation responses to different boundary markers in melodies H2, N1 & S1.

	S1 quaver rest	H2 dotted crotchet	S2 minim	N2 minim	N2 dotted minim
H2 dotted crotchet	$p > 0.09$	***	***	***	***
S2 minim	$F(1,30)=10.39$, $p < 0.005$	$F(1,35)=7.68$, $p < 0.01$	***	***	***
N2 minim	$F(1,16)=14.67$, $p < 0.005$	$p > 0.07$	$F(1,33)=4.92$, $p < 0.05$	***	***
N2 dotted minim	$F(1,16)=18.04$, $p < 0.005$	$F(1,21)=16.12$, $p < 0.005$	$p > 0.06$	$F(1,20)=18.66$, $p < 0.001$	***
S1 dotted minim	$F(1,18)=31.68$, $p < 0.001$	$F(1,23)=26.64$, $p < 0.001$	$p > 0.20$	$F(1,21)=2.02$, $p < 0.005$	$p > 0.10$

Summary of speed of segmentation results

There were significant effects of note duration on speed of segmentation response as revealed by the specific comparisons of N1 minim and N1 dotted minim (same melody with different durations). Segmentation responses occurred nearer the boundary note offset for boundary notes of shorter duration. There were no significant effects of melody type, as was demonstrated by the comparison of S1 dotted minim with N1 dotted minim (different melodies with same duration). However, there were significant differences between minims from S2 and minims from N1 with faster responding made in melody N1 compared with S2.

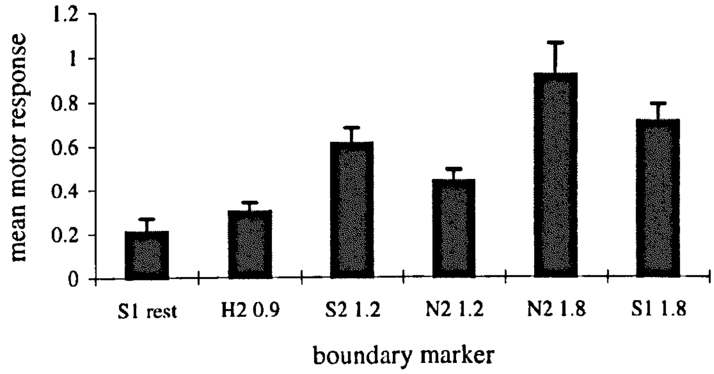


Figure 8. Motor response as a function of note length in melodies S1, H2 & N1.
Mean responses were taken for each of the major boundary markers. The y-axis displays mean response times (in seconds) *before* the boundary marker offset (i.e. before the end of the note). Thus 0.60 seconds on the y-axis indicates responses were made one crotchet before the boundary note ended, 1.2 seconds indicates two crotchets before the boundary note ended and so on. The x-axis displays the duration of the boundary notes, in seconds, in each of the three melodies. The rest in S1 is 0.30 seconds in duration.

Discussion

The main point of interest lay in identifying where segmentation responses occurred in each melody and explaining and accounting for response trends in relation to the underlying structural features of the music. The main finding is that listeners consistently segment melodies into a number of smaller units. The temporal durations of the segmented units, displayed in Figures 1 - 7, above, ranged from 2.4 seconds to 5.4 seconds, depending on the particular melody. Fraisse (1964) argues that listeners group successive events from temporal sequences into a series of time-spans. It is interesting that the four to five second time limit he places on such time-spans is reflected by the length of units into which listeners segment the melodies. Similarly, Aiello (1994) argues that we have a limited capacity for remembering the music we process and in order to process the sounds we hear most effectively we need to segment them along a temporal continuum. She offers no temporal constraint on such segments, however, and merely acknowledges that phrases have a finite time span.

The results of the present experiment revealed a powerful effect of note length on perceived boundary position within melodies H2, N1, S1 and S2, with longer notes very clearly allowing listeners to perceptually differentiate between groups in the melodies. The results of *unstructured* melody S2 raise some interesting points. Even though the melody made little musical sense (it was the first two phrases of melody S2 rearranged), there were consistent segmentation responses which split the melody into two main segments of 4.8 and 3.9 seconds in length. The segmentation responses occurred in relation to two of the longest notes in the melody. This finding is in line with the strong effect of note length on segmentation responses found in normal melodies. The first point of perceived segmentation in *unstructured* melody S2 occurred between a large pitch leap from D5 to F4, with D5 perceived as the main boundary marker. This is in line with suggestions by Thorpe & Trehub (1989) that relatively sharp changes in stimulus parameters guide the listeners segmentation of stimulus sequences into chunks which in turn provide a basis for further cognitive processing. It may be that auditory information is segmented into smaller units regardless of whether or not it actually makes musical sense to the listener. Salient features in a sequence, such as sharp changes in pitch or relatively long note durations, seem to be enough to elicit a segmentation response in the majority of listeners. Additional support for the effect of longer notes as structural markers comes from responses to melody S3. This melody, which is constructed from notes of the same duration (all crotchets), was not consistently segmented by listeners. This suggests that segmentation responses are determined primarily by salient structural markers such as note duration. The other structural feature to elicit boundary judgements from 79.2 % of listeners was the quaver rest in melody S1. It may be that the pause induced by the 0.30 second quaver rest served as a perceptual break in the melodic flow and caused the perception of a group up until the rest and another which began after the quaver rest. This finding is in line with Deliege (1987) who reports that gaps in the music are a fundamental element in grouping behaviour. The majority of non-musician participants in Deliege's (1987) study segmented musical sequences as soon as a prolonged sound occurred. This demonstrates the effect of note length as a perceptual boundary marker.

The button-press response task used in the present study raises some interesting issues. Given that listeners were required to respond to boundaries in the melodic sequences during their listening experience, it is a natural assumption that responses would be made *after* the occurrence of a boundary (Clarke & Krumhansl, 1990). Previous experiments got around this problem by allowing the listener to adjust their segmentation responses (displayed on a musical score) if they so desired. There are, however, problems of visual input biasing responses in this method (as has been discussed in the introduction). Indeed, this was one of the main reasons for utilising a segmentation task without visual information about the musical stimulus in the present study. In support of the task used in the present experiment, it should be pointed out that the majority of responses were made before the boundary offset in melodies H2, N1, S1 & S2, as is evident from Figures 2, 3, 4, 5 and 6, above. As these Figures show, there is a trend for segmentation responses to occur earlier than the boundary note offset. The majority of segmentation responses to perceived boundaries in the melodies were made *on* the sounded note rather than just *before* or just *after* it. Although there are problems associated with localising the reactive responses made in this kind of segmentation task (Clarke & Krumhansl, 1990), the present results suggest that as soon as the boundary notes sounded listeners begin making their segmentation response. Furthermore, there is evidence to suggest that listeners modified the speed of their segmentation responses in relation to the particular boundary note duration. If the phrase boundary falls on a dotted crotchet (duration of dotted crotchet is 0.90 seconds), the mean segmentation response is an average of 0.22 seconds faster than the mean response to a phrase boundary which falls on a minim (duration of minim is 1.20 seconds). Mean segmentation responses to a minim boundary are in turn an average of 0.38 seconds faster than responses to a dotted minim boundary (duration of dotted minim is 1.80 seconds). It seems that the problem of localising where listeners perceived the major boundary to be is reduced by the way in which listeners adapt the speed of their segmentation responses to the particular structural boundary. Not only do listeners generally respond before the boundary note offset, they modify the speed of their responses in relation to the boundary note duration. This suggests that listeners are aware of the particular note in the melody excerpt they want to make their segmentation response on and are aware of the length of time available to make that response. Thus we can be relatively confident that the notes on which the majority of the listeners responded are those notes which are perceived as the major boundary markers.

The tendency for listeners to respond on or as the perceived boundary note (or final note in the phrase) is sounding implies that, even in a straightforward segmentation response task, listeners are keen to avoid making responses which cross between two different groups. The present segmentation results show that as listeners make their segmentation response to indicate different groups or units in the melody, they make this response before a new group has begun (i.e. they respond *on* the boundary note). They do so even though they were instructed to make their responses once they thought a unit of meaning had occurred. This reflects the tendency of listeners to maintain the perceptual integrity of different groups within a melody (Tan et al, 1981).

It is interesting that although the melodic sequences lacked any of the timing deviations reported in previous segmentation experiments using real performances, such as temporal slowing at phrase boundaries (Palmer, 1996), listeners were fairly consistent in their segmentation patterns. It has been argued by Deliege (1987) that note durations are more often used by listeners to determine points of perceived segmentation than expressive features such as slurs. In addition to this Krumhansl (1996) argues that performance cues have little effect upon the listeners perceptual responses on a segmentation task and suggests that pitch and durational factors in the music effect segmentation responses independently of performance nuances. It may well be, as Repp (1992) suggests, that expressive deviations used in real performances of music are not essential to the listener's understanding of the music. Such timing deviations may exist to aid the structural interpretation of the performance, but it is likely that listeners form their mental structures from what ever available cues exist in the music. He suggests that if the structural features of the melody excerpts are interpretable within the realms of previous music experience, the absence of expressive performance variations should not impair the listeners understanding. Features of the musical stimuli will be enough to elicit normal perceptual responses. It seems to be the case that listeners are interpreting the present Gaelic melodies without expressive performance variations.

The conclusions drawn from the present experiment are necessarily confined the stimuli used, namely unaccompanied Gaelic melodies. Although no generalisations can be made towards large scale multilayered pieces of music, the results are of the utmost importance for a consideration of the factors which form the basis of the listeners preferred segmentation responses in unaccompanied melodies. As Fodor & Bever (1965) point out, there are a number of experimental techniques which we can use to determine the segmentation of a complex percept. The most straightforward of these is to ask the listener directly for their preferred segmentation of a sequence of events as was done in the present experiment. Having established the preferred perceptual units of these types of melody using this simple method and considered what factors play the most important role in determining these perceptual boundaries, we can now examine how these different boundaries may bias listeners' responses in other more subtle perceptual tasks. One such task which has been widely used to examine processing units within speech and, to a lesser extent, in music is the click localisation task to which this discussion now turns.

Experiment II Click localisation in Gaelic melodies

Background

Click localisation is a useful technique for examining perceptual processing units within speech (Fodor & Bever, 1965; Ladefoged & Broadbent, 1960) and music (Stoffer, 1985). The basic task involves the listener attending to an auditory stimulus in which a superimposed click is positioned and subsequently indicating when the click was perceived to have occurred. A basic assumption of these experiments is that the pattern of errors which listeners make in identifying where the click occurred in a sequence give us an insight into the way in which listeners are processing the auditory information (Fodor & Bever, 1965; Reber & Anderson, 1970). An examination of the perceived location of clicks will reveal errors in direction and magnitude and emphasise positions in the sequence which attract click migration. This highlights the occurrence of information units which listeners may use in processing auditory sequences.

The general approach has been to relate the perceptual responses of the listener to the grammatical structure of the stimuli in a systematic manner (Fodor & Bever, 1965). The main focus with musical stimuli has been to explain or relate click localisation errors to the phrase structure of the music (Gregory, 1978; Sloboda & Gregory, 1980; Stoffer, 1985). There are a few researchers who have considered migration results in relation to attentional factors within a temporal framework (Reber & Anderson, 1970; Reber, 1973) but on the whole this type of approach remains neglected.

Results of click localisation experiments with speech and music generally show that listeners tend to perceptually migrate the click towards the nearest phrase boundary. Basically, clicks positioned before a boundary will be perceived as occurring later than their correct position and clicks positioned after the phrase boundary will be heard as occurring before their actual position. This has been widely interpreted as demonstrating the listeners tendency to process the incoming auditory stimuli as discrete units of information. Click migration results are reported to reveal the manner in which the listener preserves these information units. The listener prefers to hear superimposed clicks at positions of naturally occurring breaks in the stimuli, i.e. at the end of an information unit, rather than within a group of tones which do not form complete information unit. As discussed above, (Thorpe, Trehub, Morrongiello & Bull, 1988; Fraisse, 1982, in the Segmentation Experiment), these perceptual breaks are most likely to occur between phrases rather than within phrases, as shown by preferences for pauses between groups of tones rather than within groups of tones.

Click migration and music

Click migration in music has been demonstrated in short artificial sequences of tones (Gregory, 1978) and rhythmically simple two-phrase melodies (Sloboda & Gregory, 1980), using a visual task in which listeners mark the perceived click on a musical score. Methodological criticisms raised by Stoffer (1985) include the confounding effect of visual segmentation of the sequences and the use of artificial melodies which lack 'real' phrase indicators. These issues and others will be discussed in turn.

The stimuli used by Gregory (1978) were six note sequences which were visually grouped into two lots of triplets or three sets of two quavers. Problems with using this type of stimuli for the purposes of click experiments are two-fold. The length is not representative of that of most melodies. Played at two different speeds the overall duration of the sequences was 1.1 or 2.2 seconds long. As pointed out by Sloboda & Gregory (1980), when listeners attend to this type of short stimuli they may utilise processes which differ from those used when listening to longer pieces of music. In addition to this, all notes were of equal duration which is uncommon for most of the melodies which listeners perceive. Rhythmic structure is an important part of any musical sequence and such a simplistic equitemporal arrangement may also alter the type of processing which listeners engage in. It has been suggested that melodies of at least two phrases should be used in click experiments (Sloboda & Gregory, 1980). Another problem with the task demands of Gregory's (1978) experiment arose because of the enforced segmentation of notes which listeners were instructed engage in. The alternatives: two groups of three triplets or three groups of two quavers, may not be the natural groupings, if any, which listeners would use for these six note sequences. The migration effects, which were reported at these points of experimenter-defined segmentation, do not add much to our understanding of how listeners process information within phrases. The second main problem with the experiment was the visual task. The listener had to think of the sequence as either two or three groups of notes (they were given a score with the corresponding arrangement of grouped notes) and point to the note or pause in the sequence where the click occurred. The 'phrase' migration results reported by Gregory (1978) are necessarily tied up with the visual segmentation of the two different scores and as a result cannot be interpreted solely as auditory musical phrase boundary effects.

Sloboda & Gregory (1980) used longer two-phrase sequences of 20 or 21 notes in length, or 4.6 seconds duration, with a similar experimental procedure. Subjects were given a copy of the visual score which they viewed whilst listening to the melody. On hearing the click, they marked the position at which they thought the click had occurred onto the musical score. The melodies were equitemporal (quaver notes) apart from the last note of each phrase which was a crotchet. The last notes in the phrase were altered so that the melody would end on one of the following four ways:

1. A crotchet note with cadence
2. A crotchet note, pitches altered to eliminate harmonic sense and retain the melodic contour in 1.
3. Two quaver notes which preserved the harmonic sense of the melody in 1.
4. Two quaver notes, pitches altered to eliminate harmonic sense and retain the melodic contour in 1.

This experiment failed to address the potential effect of visual segmentation of the melodies with phrases which ended on crotchet notes. It has been noted that phrases are often separated from each other by long time intervals at the phrase boundary, such time intervals are reflected in the music notation with longer spaces occurring with longer durations (Sloboda, 1977). As a result the spaces in music notation may be used as phrase boundary markers. The potential confounding effects of visual segmentation should be taken seriously. This is especially important in a study which used crotchet notes as boundary markers in a melody of quaver notes. The crotchets were given double the spacing which the quavers were given.

Sloboda & Gregory (1980) interpret their migration effects as evidence that the listeners' cognitive representation of melodies are in line with the phrase structure formally identified by professional musicians. However, they also acknowledge that the visually salient crotchet note may have been used as a reference point with which to track the melody. They propose that their experimental effects are only partly perceptual and propose that further research is required in order to examine the confounding auditory and visual effects on migration responses.

Stoffer (1985) used click migration to examine whether there is a structural correspondence between the listeners' cognitive representation of music and the phrase structure of such music. Despite pointing out the methodological hazards of the type of visual task used by Gregory (1978) and Sloboda & Gregory (1980), Stoffer (1985) employed a similar visual task in his localisation experiments. The main difference was that listeners were not allowed to see the musical score until they had heard the click. Listeners marked their response on the musical score which does not completely overcome the problem of confounding visual information. Rather, it introduces the same problem at a slightly later stage in the experiment.

Response bias?

It is argued that the click localisation method is subject to response biases such as a greater number of early migration responses when attention is focused on the click sound and the tendency to locate 'sub-threshold' clicks at major phrase boundary positions in speech (Reber & Anderson, 1970; Reber, 1973). Reber (1973) argues that it is futile to explain migration effects in terms of structural features in the stimuli, without considering the underlying response distribution for that stimuli. He maintains that the click migration effect around a phrase boundary is called into question if listeners consistently respond there when a click is not actually heard. Reber (1973) proposes that any interpretation of the results should take account of where the listener is most likely to consider a click as occurring. Others argue that the click migration effect cannot be explained in terms of response bias (Dalrymple-Alford, 1976) and it is suggested that click effects are entirely perceptual in nature and result from the way listeners segment a sequence during processing (Fodor, Bever & Garrett, 1974). Gregory (1978) argues that the response bias reported by Reber (1973) may actually demonstrate the reality of segmental groups in the stimuli. It can be argued that the tendency of listeners to locate imperceptible clicks at phrase boundaries supports the idea that listeners prefer to maintain discrete units of information when processing incoming sequences. Stoffer (1985) proposes that in order to minimalise response bias, the listener must focus on the musical structure rather than on the click.

Problems with the dichotic listening task

Another methodological problem is the dichotic listening procedure which is predominantly used for click migration experiments. The click is presented to one ear while the musical stimuli is presented to the alternate ear. An effect of ear of presentation can occur using this method for both click perception and melody recognition (Kimura, 1964) with faster processing of clicks or musical stimuli which are presented to the left ear. In addition to this, the nature of the dichotic task, one stimuli in one ear (e.g. click) and the other in the alternate ear (e.g. music) encourages the listener to dissociate the music and the click and hear them from two distinct sources. This in turn exacerbates the problem of listeners directing their attention to one sound source over the other when performing the experiment. Although some click localisation studies control for this possible confounding factor by presenting clicks to left and right ears with musical stimuli in the alternate ear (Gregory, 1978; Sloboda & Gregory, 1980), others do not (Stoffer, 1985).

Gregory (1978) found that clicks were perceived as occurring earlier if they arrived at the left ear (music in the right ear) than the other way about. Sloboda & Gregory (1980) found that the clicks were perceived as occurring later than their correct position on 75 % of trials when presented to the right ear with music to the left ear. Stoffer (1985) presented clicks to the right ear only and melody to the left ear in all his experimental trials. The rationale behind this was to use the presentation which had produced the maximum click displacement in previous studies. Given that early click displacement could be enhanced as a result of hemispheric differences, left ear dominances, or directed attending to the click sound over the music, clicks should have been presented to alternate ears as a minimal procedural requirement in the experiment. The only real solution to these methodological problems is to completely abandon the dichotic task in click localisation experiments.

Where do listeners segment the sequences?

Another criticism is that click localisation studies have neglected to consider where the listener would naturally parse the experimental sequences. Rather, they have focused on relating any migration effects obtained to formal descriptions of the phrase structure of sequences presented to listeners. Listeners in Stoffer's (1985) study performed a detailed 12 hour training schedule in which they were taught to recognise phrase structure markers of the experimental melodies which were artificially generated from a transformational syntax of German folk songs. Given the highly artificial manner in which listeners were exposed to these melodies, there is a serious limitation in considering the results to be informative about the processing strategies employed when listening to real music.

Sloboda & Gregory (1980) aimed to explore whether the musical phrase (and by phrase they meant that which trained musicians identify by consensus) can be shown to be psychologically real using a click migration experiment. In other words, do click migration effects reveal the listeners representation of a phrase to be the same unit which is formally defined by trained musicians? Phrase boundaries formally specified by trained musicians *may* be the most likely perceptual segmentation points in a given sequence. Similarly, the syntactical analysis used by Stoffer (1985) can establish the phrase structure of the experimental material which *may* reflect the representational units used by the listener. However, these assumptions have not been directly tested which suggests that there is a need to empirically define perceptual phrases.

The present study

The most straightforward method for doing this is to ask the listener directly for their preferred segmentation of a sequence of events. As discussed in Experiment I, this was the main aim of the segmentation research. Having established that there are preferred perceptual units in the Gaelic melodies, a set of experimental conditions were performed to examine a novel click localisation task in relation to the perceptual boundaries of melodies identified in the segmentation experiment. The main differences in the present task compared with previous click localisation studies are as follows:

1. The *binaural* presentation of real musical stimuli (unaccompanied Gaelic melodies) as opposed to the dichotic technique typically employed to examine responses to artificial melodies. The listener heard identical stimuli (melody excerpt containing a click) in both ears at the same time.
2. The altered non-visual task which avoided visual segmentation effects by employing a non-visual motor and memory task in comparison to previous studies which used the musical score of the melodies. The present task can be performed by non-musicians.
3. An examination of migration effects around boundaries which have been objectively identified in a segmentation experiment rather than testing listeners representational units in relation to experimenter-defined phrases.

The experimental task is detailed below in the methods section. Seven experimental conditions were carried out to examine localisation responses under a variety of conditions. The first condition reports the findings from a pilot study. Given the encouraging findings from Condition 1, Condition 2 extended the use of the novel localisation task to a larger number of participants with a wider range of click positions. When utilising a novel task it is important to ensure that, for example, listeners are attending to the musical structure as directed by the task instructions and that features of the task do not contribute directly to the experimental results. Condition 3 was run to check that listeners were attending to the musical structure as directed. Condition 4 was a check on the localisation task itself. A number of interesting findings were generated by the localisation studies. Conditions 5 and 6 aimed to account for these results in a specific way. Finally, Condition 7 used the localisation task with unstructured musical stimuli. Responses to this stimuli are considered in relation to musical stimuli.

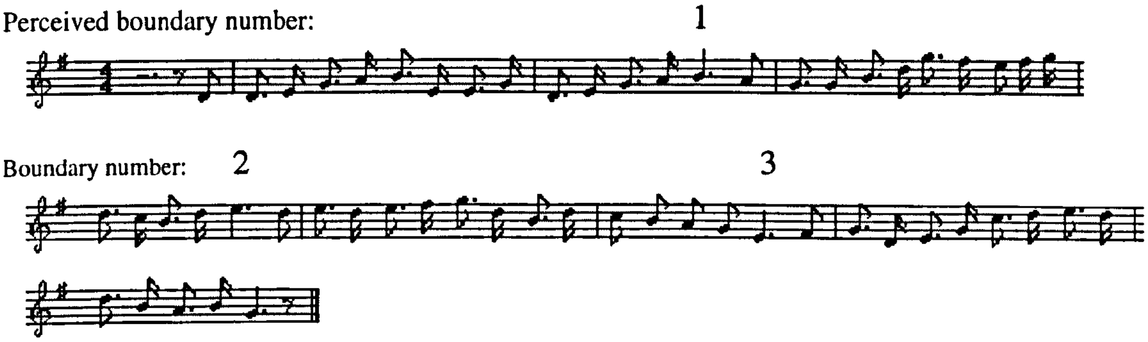
Condition 1 Click localisation in Gaelic melodies - the pilot study

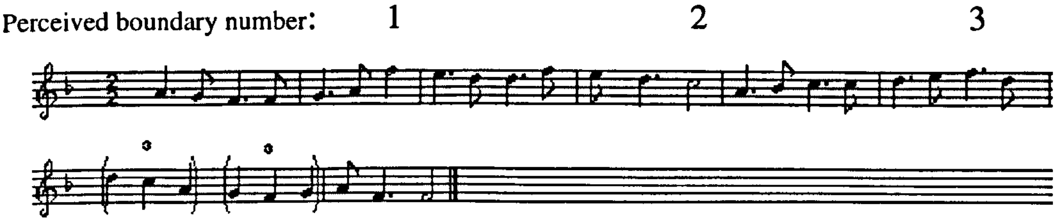
Method

Subjects The subjects were 19 psychology postgraduates, 8 females, 11 males, aged between 24 and 35 years.

Stimuli/Apparatus The stimuli consisted of excerpts from melodies H2 and S2 shown in Figure A, below. The excerpts were produced in exactly the same way as in the Segmentation Experiment (Experiment I), above. The additional click sounds were created by inserting a 50 msec amplified noise burst on a note onset or in the middle of a note. Each excerpt contained *one* click from the chosen range of possible click positions. These were recorded onto audio cassette tape and were followed immediately by a copy of the excerpt *without* the click sound. Stimuli were played on a SONY F442E stereo cassette deck at comfortable sound listening levels. Four different recordings were made, randomising the click position to control for order effects. As in the Segmentation Experiment, listeners made their response using a hand-held button press. Their responses were detected by an electronic switch and recorded with a sound edit package on a Macintosh Centris 650.

Figure A Melodies used in click localisation condition 1

Melody H2

Melody S2



Procedure The following two click detection experiments were performed prior to the main experiment:

1. Click detection in response to isolated clicks presented without warning. This was to examine speed of responses to the click sound outwith the musical context. Listeners were asked to press the button as soon as they heard the click sound.
2. Click detection in response to a click embedded within an isolated note. The notes were selected from melodies S2 and H2 and were presented individually. Listeners were asked to press the button as soon as they heard the click sound and were also required to indicate if the click occurred as the note sounded or at the end of the note.

These conditions were carried out to ensure that any click detection effects obtained in the melodies could not be attributed to an increased difficulty of detecting the click sound within a tone.

For the main localisation experiment, listeners performed with excerpts from one of the melodies only. Ten listeners performed with melody H2 and nine with melody S2. They heard two excerpts in close succession. The first excerpt contained one click placed on a note onset or in the middle of a note in the melody. In response to hearing the click listeners pressed the button. The second excerpt, a copy of the first *without* the click but identical in every other respect, was then heard and listeners were required to press the button when the click was perceived to have occurred. The listener was instructed to attend to the melody primarily, rather than the click, in the click detection task in order that they could perform the subsequent localisation task as accurately as possible. The experiment ran as follows:

1. Excerpt 1: click is placed somewhere on a note onset or mid-note position. The listener performs a click detection task by pressing the switch when the click is perceived.
2. Excerpt 2: excerpt 1 without the click. The listener performs a memory task by pressing the switch where they think the click occurred in excerpt 1.

Practise trials were run using excerpts from two similar Gaelic melodies to enable the listener to become accustomed to the task. Listeners were tested individually with the experimenter. Experimental trials lasted between 30 and 40 minutes.

Results

The click localisation results (where listeners think the click occurred) will be presented first followed by the click detection results (reaction times to the click sound).

Click localisation

The main interest lay in examining localisation responses in relation to the three perceived boundaries in each melody. Are localisation errors late or early? And if so, by how much? This was determined by subtracting the *objective* click position (where the click actually occurred in the experimental stimuli) from the *subjective* position of the click (where listeners perceive the click as occurring). This gave an error of response measure (in seconds) for every different click position for each listener. Positive errors indicate *late* localisation responses (listeners perceived the click as occurring *later* than its actual position) and negative errors indicate *early* localisation responses (listeners perceived the click as occurring *earlier* that its actual position). The distinction between subjective and objective click position is an important one. From now on, click position will refer to the position at which the click occurred in the experimental stimuli (the objective position). Click localisation will refer to the listeners subjective response to the click which is positioned in the stimuli. Click migration will also refer to the subjective displacement of the click from its actual position in the experimental stimuli.

Data was grouped into localisation responses to clicks positioned *before* and *after* main perceived boundaries 1, 2 & 3, for the two melodies. This was done by taking the mean localisation response to clicks objectively positioned within two crotchet beats before and two crotchet beats after each boundary. A comparison of click localisation errors for clicks positioned on note onsets with clicks positioned on the corresponding mid-notes was carried out. A repeated-measures ANOVA was performed separately on responses to clicks positioned before and clicks positioned after boundaries 1, 2 & 3. Position on note was the within-subjects factor (click on note-onset or click on mid-note) and melody was the between-subjects factor (H2 and S2). For clicks positioned before the boundaries, there were no significant within-subjects effects, $p > 0.10$, no effect of melody, $p > 0.20$ and no interaction, $p > 0.50$. The results were the same for clicks positioned after the boundaries, with no within-subjects effect, $p > 0.60$, no effect of melody, $p > 0.20$ and no interaction $p > 0.70$. To summarise, localisation errors for clicks positioned on note-onsets were similar to localisation errors for clicks positioned in the middle of these notes. This effect was the same for both melodies. The similarity of localisation response to clicks positioned on and clicks positioned in the middle of a note occurred for notes before and after the three main boundaries. Means are displayed in Tables 1 & 2, Appendix A, page 88.

A repeated measures ANOVA was performed to compare the localisation response to clicks positioned before and clicks positioned after boundaries 1, 2 & 3. Boundary position (before or after) was the within subjects factor and melody (H2 or S2) was the between subjects factor. This revealed a significant main effect of click position, $F(1, 17) = 14.20, p < 0.005$. Clicks positioned *before* boundaries were perceived as occurring *later* than their actual position (0.07 seconds). Clicks positioned *after* boundaries were perceived as occurring *earlier* than their actual position (0.15 seconds). Means for boundaries 1, 2 & 3 in each melody are displayed individually in Figures 1 and 2, below. There was no effect of melody type, $p > 0.20$ and no interaction, $p > 0.60$. Thus the effect of click position was robust for both melodies.

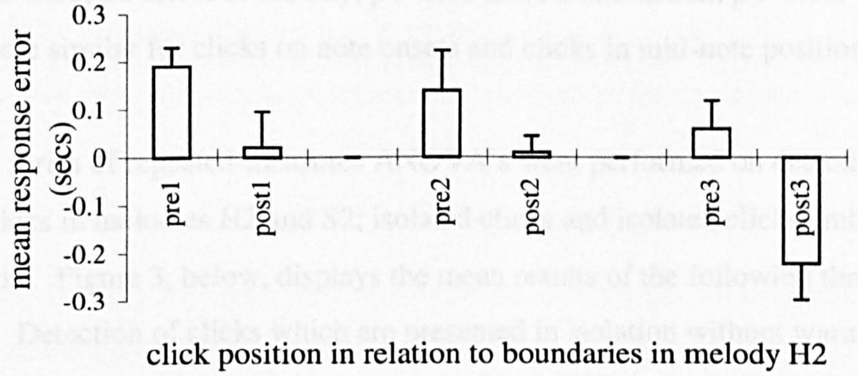


Figure 1 Localisation error of clicks positioned before and after main boundaries in melody H2. The x-axis corresponds to clicks positioned before (p*re*) and after (*post*) the main boundaries identified in the Segmentation experiment (Experiment I). The number corresponds to the perceived boundary number. The y-axis represents the mean response errors. Zero corresponds to a correct response (click is perceived in the position in which it occurred), a positive score is a late migration (click is perceived as occurring later than its correct position) and a negative score is an early migration (click is perceived as occurring earlier than its correct position).

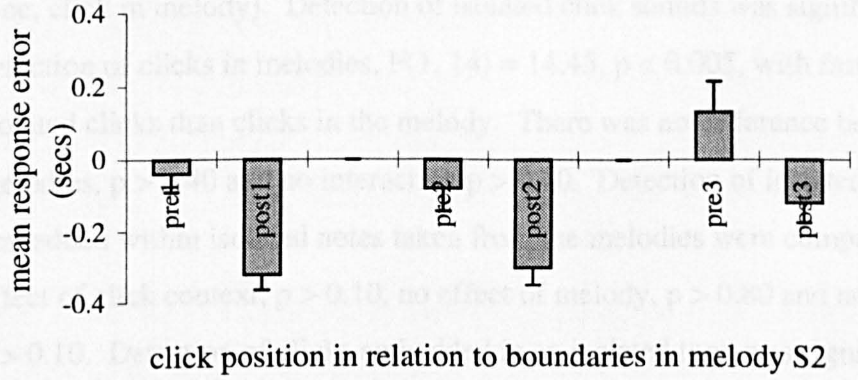


Figure 2 Localisation error of clicks positioned before and after main boundaries in melody S2. As described in Figure 1, above, the x-axis displays the click position in relation to the three perceived boundaries and the y-axis shows the localisation response to clicks before and after each boundary.

Click detection results

Detection of clicks in the two melodies were analysed. The time between actual click occurrence and button-press response in the initial detection task was used as the reaction time measurement. A repeated-measures ANOVA was performed on the click detection data for clicks placed on note-onsets and clicks placed in the middle of corresponding notes. The ANOVA revealed that there was no significant difference between detection times of clicks on note-onsets and clicks in mid-note positions, $p > 0.10$, no effect of melody, $p > 0.60$ and no interaction, $p > 0.09$. Detection times were similar for clicks on note onsets and clicks in mid-note positions in both melodies.

A series of repeated-measures ANOVA's were performed on detection responses to: clicks in melodies H2 and S2; isolated clicks and isolated clicks embedded within a note. Figure 3, below, displays the mean results of the following three groups:

1. Detection of clicks which are presented in isolation without warning
2. Detection of clicks which are embedded within or at the end of a note presented in isolation without warning
3. Detection of clicks which are embedded within a melody.

Sixteen listeners performed the click alone condition, eight of the listeners performed with melody H2 and the isolated notes from melody H2 and eight listeners performed with melody S2 and the isolated notes from melody S2. Melody is the between-subjects factor (H2 and S2) and type of stimuli is the within-subjects factor (click alone, click in tone, click in melody). Detection of isolated click sounds was significantly different to detection of clicks in melodies, $F(1, 14) = 14.45$, $p < 0.005$, with faster detection of isolated clicks than clicks in the melody. There was no difference between the two melodies, $p > 0.40$ and no interaction $p > 0.40$. Detection of isolated clicks and clicks embedded within isolated notes taken from the melodies were compared. There was no effect of click context, $p > 0.10$, no effect of melody, $p > 0.80$ and no interaction $p > 0.10$. Detection of clicks embedded in an isolated tone was significantly different to detection of clicks within a melody, $F(1, 14) = 10.98$, $p < 0.005$, with faster detection of clicks embedded in isolated tones. Again, there was no effect of melody type, $p > 0.80$ and no interaction $p > 0.20$. The results obtained are similar for both melodies. Means are also displayed in Tables 3 & 4, Appendix A, page 88.

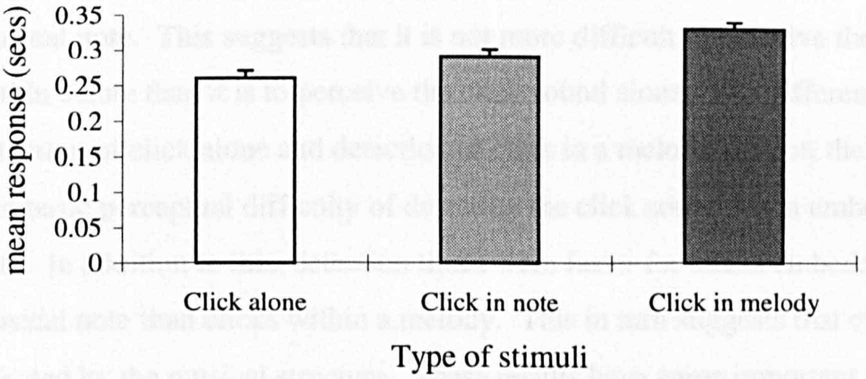


Figure 3 Mean click detection times as a function of stimuli context.
The x-axis describes the stimuli used in the detection experiment. The y-axis displays the mean detection times (in seconds). Detection data is averaged across all note positions for both melodies.

Discussion

The present pilot study generated some useful findings. The comparison of mid-note with note-onset click positions was carried out to resolve the ambiguity surrounding the positioning of clicks on experimental material and to ascertain whether click position on a given note has an effect on detection and localisation performance. The results of the present experiment show that localisation errors were similar for clicks which were positioned on a note and clicks positioned in the middle of a note. Click detection times were also similar for clicks which were positioned on note onset and mid-note positions. This suggests that clicks can be positioned on a note or in the middle of a note without affecting click localisation or click detection performance. However, it has been suggested that any click sound which is asynchronous with a note onset will tend to be heard as occurring after that note (Sloboda & Gregory, 1980). Considering the results of the present experiment in conjunction with such suggestions, the stimuli in Conditions 2 - 7, below, contained clicks which were positioned on note onsets only.

Click detection was significantly faster (an average of 0.07 seconds faster) for isolated clicks with no warning in comparison to clicks which were embedded within a melody. It was important to ensure that this effect was not due to an increased difficulty of detecting a click embedded within a musical note compared with detection of a click on its own. Comparisons were made between detection of clicks alone and detection of clicks embedded within a musical note. The results demonstrated that there were no differences between detection of clicks alone and detection of clicks embedded within a musical note. This suggests that it is not more difficult to perceive the click sound within a note than it is to perceive the click sound alone. The difference between detection of click alone and detection of click in a melody are not, therefore, the result of a basic perceptual difficulty of detecting the click sound when embedded within a note. In addition to this, detection times were faster for clicks embedded within a musical note than clicks within a melody. This in turn suggests that click detection is affected by the musical structure. These results have some important implications. First of all, the results suggest that listeners are attending to the melody primarily rather than listening to the click in isolation. This satisfies one of the main requirements for controlling response bias. It is suggested that the click localisation task is only useful if the listeners attention is focused on the musical structure rather than on the click (Stoffer, 1985). The results also imply that click detection response patterns (e.g. significantly slower/faster click detection responses to certain positions in the melody) can be used to highlight areas of high and low information in the musical stimuli. The link between click detection and information in the melody can be explained more clearly by considering previous click localisation research. Ladefoged & Broadbent (1960) suggest that the speed at which listeners respond to a given stimulus (e.g. a click sound within the context of a melody) is related to how much attention they are giving to that stimuli. Attention will be directed towards the melody if there is a high probability that an information bearing event will occur and click detection times will be slower at these points. This suggests that positions in the melody which yield low information will yield faster click detection responses. The click detection data can therefore be used to highlight the pattern of information in the musical structure and may highlight the expectancies of the listener in the course of their music listening.

The present click localisation results showed migration effects, around the main perceived boundary 3 in melodies S2 and H2, which are in line with previous click localisation experiments (Gregory, 1978; Sloboda & Gregory, 1980; Stoffer, 1985). Clicks positioned before the boundary were perceived as occurring later than their actual positions and clicks positioned after the boundary were perceived as occurring earlier than their actual position. It is important to note that these results were obtained in real melodies using a task which eliminated the confounding effect of visual segmentation of the musical score. In addition to this, problems of the widely used dichotic listening task were removed by presenting the experimental stimuli binaurally. Thus we can be fairly confident that this particular click localisation task is useful for examining the way in which information in melodies is processed. Condition 2 was a follow-up study which was carried out with a larger participant sample who performed the task over a wider range of click positions in different melodies. In addition to this, the effect of familiarity with the melodies on task performance was examined.

Condition 2 Click localisation in Gaelic melodies

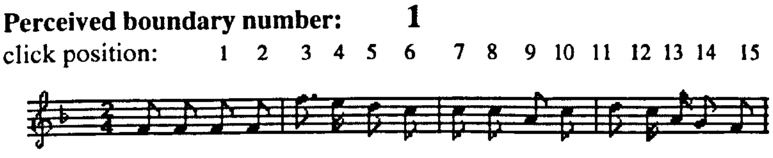
Method

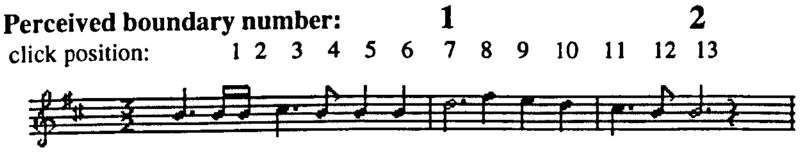
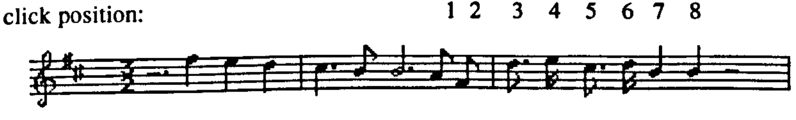
Subjects The subjects were 33 psychology undergraduates, 20 females, 13 males, aged between 17 and 47 years. Participation in the experiment was in part fulfilment of a psychology course requirement.

Stimuli/Apparatus The stimuli consisted of excerpts of the first two perceived boundaries identified in melodies H1 and S1 in the Segmentation Experiment (Experiment I, above) and excerpts of phrases 2-3 and 3-4 in melody S1. The excerpts are shown in Figure B and described in Table 1, below.

Figure B Melody excerpts used in click localisation Condition 2

Melody H1



Melody S1: Phrase 1 -2

Melody S1: Phrase 2 - 3

Melody S1: Phrase 3 - 4

click position:

Table 1 Basic characteristics of the four experimental excerpts

Melody	Duration (secs)	Number of notes	Number of click positions	Range of clicks
S1	10.2	14	13	note 2 - note 14
S1	7.8	14	8	note 15 - note 21
S1	7.8	16	8	note 22 - note 29
H1	4.8	17	15	note 2 - note 17

The excerpts were produced in exactly the same way as condition 1, this time the clicks were recorded onto note-onsets alone and the listener heard the same excerpt for each of the possible click positions within a phrase. For melody S1, this resulted in 13 excerpts for phrase 1-2, 8 excerpts for phrase 2-3 and 8 excerpts for phrase 3-4 and 15 excerpts for phrase 1-2 in melody H1. Within each excerpt the only difference was click position, all other factors such as overall duration and loudness were identical. Each excerpt containing a click was recorded onto audio cassette tape and followed by a copy of the excerpt without the click but identical in every other way. These were played on a SONY F442E stereo cassette deck at comfortable sound listening levels. Listeners responses were detected and recorded as in Condition 1. Ten different recordings, which varied the order of click positions, were made to control for order effects.

Procedure Listeners were split into two different groups of 19 unfamiliar listeners and 14 familiar listeners. All 33 participants performed with both melodies. The familiar listeners were given an audio cassette tape of the two melodies which they were required to play for ten minutes each day during the course of a week. Their task was to familiarise themselves with the melodies so that they could sing them back in their heads. The familiar listeners then performed exactly the same experiment as unfamiliar listeners. The experimental procedure was the same as that of Condition 1 above: a click detection task was followed by a click localisation task. Practice trials were run using excerpts from two similar Gaelic melodies to enable the listener to become accustomed to the task. Experimental trials lasted an average of 40 minutes.

Click localisation results

The mean localisation response is presented in Figures 4a and 5a, below (means in Table 1a, Appendix B, page 89). A series of repeated measures ANOVA's were performed on the data to examine the effects of click position and familiarity on click localisation performance. The between-subjects factor was familiarity and within-subjects factor was position of click in the melody. The results from each melody will be discussed in turn.

Melody H1

Familiarity

There were no significant between-subjects effects with F(1, 31) = 0.01, $p > 0.90$. Thus familiarity with the melody did not significantly alter click localisation performance.

Click position effects

There was a significant effect of click position, F(14, 434) = 11.08, $p < 0.001$ with no familiarity-click position interaction, $p > 0.70$ (means in Figure 3a, below). Click localisation performance differed significantly depending on the position of the click in the melody and this pattern of response was similar for familiar and unfamiliar listeners.

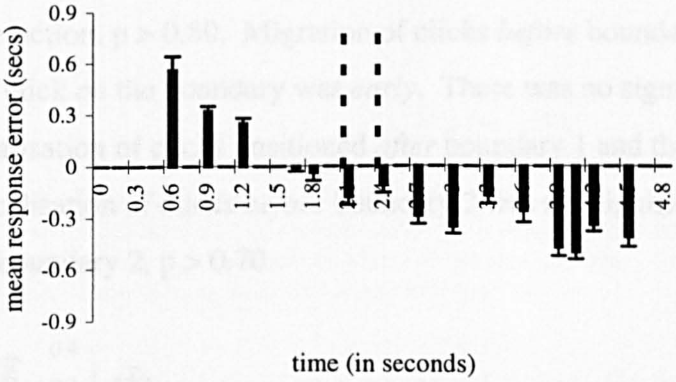


Figure 4a Click localisation response in phrase 1 of melody H1.
The x-axis represents time in seconds, each bar corresponds to a click positioned on a note in the melody. The dotted lines mark the onset and offset of boundary 1. The y-axis shows the mean localisation response in seconds. A response at zero indicates a correct response, a positive response time represents a late migration and a negative response time represents an early migration.

The main interest was in the localisation responses to clicks positioned on notes before and clicks positioned on notes after boundary 1 in the melody. The localisation data was grouped into responses to clicks positioned within 2 crotchet beats before and 2 crotchet beats after perceived boundary 1. This grouped analysis resulted in the following three separate groups for melody H1:

1. The four notes before boundary I
2. The four notes after boundary I
3. The four notes before boundary II

A repeated-measures ANOVA revealed a significant effect of click position on localisation response. Means of responses around boundary 1 are displayed in Figure 4b, below. Localisation of clicks positioned *before* boundary 1 were significantly different to localisation of clicks positioned *after* boundary 1, with $F(1, 31) = 21.74$, $p < 0.001$. Generally, migration of clicks positioned on notes *before* boundary 1 was *late* and migration of clicks positioned *after* boundary 1 was *early*. There was no significant effect of familiarity, $p > 0.21$ and no familiarity-click position interaction, $p > 0.20$. Localisation of clicks positioned on notes before and after boundary 2 were not compared because the excerpt did contain notes occurring after boundary 2. A comparison of localisation of the click *on* boundary 1 with localisation of clicks positioned *before* and *after* boundary 1 was performed. Localisation was significantly different between clicks positioned *before* boundary 1 and the click positioned *on* boundary 1, $F(1, 31) = 11.39$, $p < 0.002$, with no effect of familiarity, $p > 0.70$ and no interaction, $p > 0.80$. Migration of clicks *before* boundary 1 was *late* and migration of the click *on* the boundary was *early*. There was no significant difference between localisation of clicks positioned *after* boundary 1 and those *on* boundary 1, $p > 0.70$. Localisation of clicks *before* boundary 2 was not significantly different from the click *on* boundary 2, $p > 0.70$.

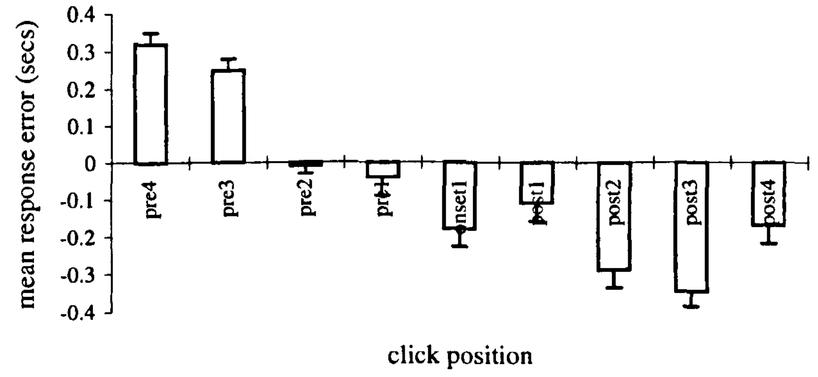


Figure 4b Magnified migration effects around boundary 1 in melody H1.
The x-axis displays the click position. Pre4 refers to four notes before boundary 1, onset1 to the boundary 1 note, post4 to four notes after boundary 1 and so on. Mean localisation response is shown on the y-axis as in Figure 4a, above.

Melody S1

A series of repeated measures ANOVA's were performed on the data. Phrase 1-2 results will be considered in detail. Mean responses are shown in Figure 5a, below and in Table 2a, Appendix B, pp 89 - 90.

Familiarity

There were no significant between-subjects effects for melody S1, $p > 0.10$. Thus familiarity with the melody did not effect localisation performance.

Click position effects

Significant within-subjects effects were found, with $F(12, 372) = 21.63$, $p < 0.001$ and a non-significant interaction $p > 0.20$. Click localisation differed significantly depending on the position of the click. The effect of click position on localisation response was similar for familiar and unfamiliar listeners.

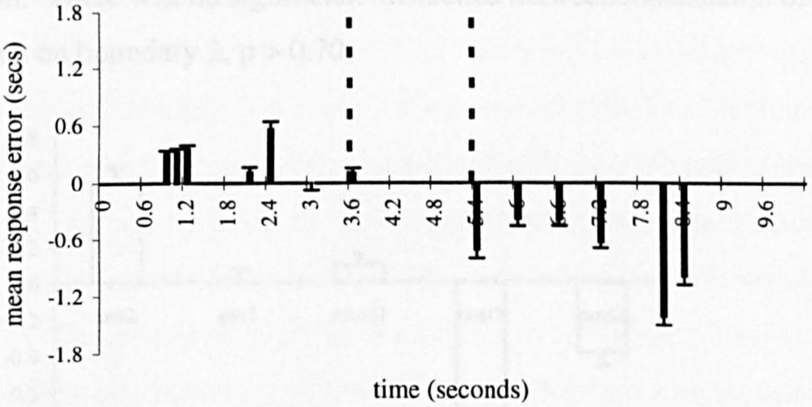


Figure 5a Click localisation response patterns in phrase 1 -2 of melody S1.
The x-axis represents time in seconds, each bar corresponds to a click positioned on a note in the melody. The dotted lines mark the onset and offset of boundary 1. The y-axis shows the mean localisation responses in seconds, as described above in Figure 4a.

The data was grouped in the same way as melody H1 for further analysis. This resulted in the following three groups:

1. The two notes before boundary I
2. The two notes after boundary I
3. The two notes before boundary II

The unfamiliar and familiar listeners data were combined due to the similarity between them. A series of repeated measures ANOVA's were performed on the grouped data with the following results. Localisation of clicks which were positioned on notes *before* boundary 1 was significantly different to localisation of clicks immediately *after* boundary 1, $F(1, 31) = 24.02$, $p < 0.001$. Clicks placed before the boundary were migrated an average of 0.28 seconds *later* than their correct position and clicks following the boundary were placed an average of 0.54 seconds *earlier* than their actual position. Means are displayed in Figure 5b, below. Localisation of clicks positioned before and after boundary 2 were not compared because the excerpt did not contain notes occurring after boundary 2. Localisation of the click positioned *on* boundary 1 was compared with those occurring directly *before* and *after* boundary 1. There was no significant difference between localisation of clicks positioned on notes *before* boundary 1 and the click *on* boundary 1, $p > 0.30$. Localisation of clicks positioned *after* boundary 1 was significantly different to localisation of the click positioned *on* boundary 1, $F(1, 31) = 10.56$, $p < 0.005$. Clicks occurring *after* the boundary were migrated an average of 0.54 seconds *earlier* than their actual position while the click positioned *on* boundary 1 was displaced an average of 0.28 seconds *later* than its actual position. There was no significant difference between localisation of clicks before and the click on boundary 2, $p > 0.70$.

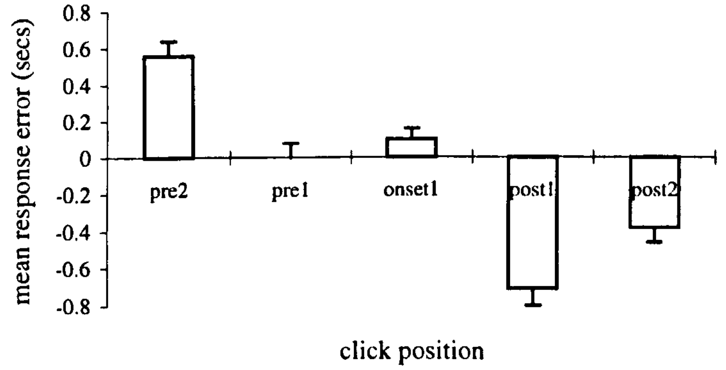


Figure 5b Magnified migration effects around boundary 1 in melody S1.
The x-axis displays click position and the y-axis shows the mean localisation response as in Figure 4b, above.

Click detection results

Click detection responses to the two melodies were analysed as above in Condition 1. A series of repeated measures ANOVA's were performed on the data. The between-subjects factor was familiarity and the within-subjects factor was click position. Means for melodies H1 and S1 are displayed in Figures 4c and 5c, below (and Tables 1b & 2b, Appendix B, pp 89 - 90).

Click position effects

The ANOVA revealed a significant overall effect of click position on click detection time for both melodies. There was a significant main effect of click position for melody H1, $F(14, 434) = 7.76$, $p < 0.001$, no significant between-subjects factor, $p > 0.80$ and no significant interaction, $p > 0.90$ level. There was a significant click position effect for melody S1, $F(28, 868) = 9.16$, $p < 0.001$. There was no between-subjects effect, $p > 0.30$ and there was no significant interaction, $p > 0.40$. Thus there were no significant effects of familiarity on click detection performance. As can be seen from Figures 4c and 5c, there is a trend for faster click detection the farther into the melody the click is positioned, especially for the first 2.5 seconds of each melody. Click responses level off after the first few seconds into each melody. A comparison of detection of the click positioned on boundary 1 with clicks positioned immediately before and after boundary 1 were carried out for both melodies. For melody H1, there were no differences between detection of clicks positioned *before* and the click positioned *on* boundary 1, $p > 0.90$. There was no difference between detection of clicks positioned *after* boundary 1 compared with the click positioned *on* boundary 1, $p > 0.10$. For melody S1, there was a significant difference between detection of clicks positioned *before* boundary 1 and the click positioned *on* boundary 1, $F(1, 31) = 5.21$, $p < 0.05$. The click positioned *on* boundary 1 were detected faster than clicks positioned *before* boundary 1. There was no difference between detection of clicks positioned *after* and the click positioned *on* boundary 1, $p > 0.10$.

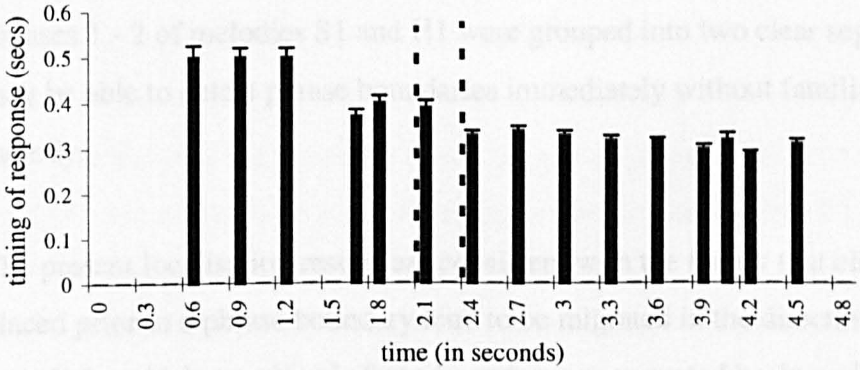


Figure 4c Click detection responses in phrase 1 of melody H1.
The x-axis represents time in seconds, each bar corresponds to a click positioned on a note in the melody. The dotted lines mark the onset and offset of boundary 1. The y-axis shows the mean response times in seconds.

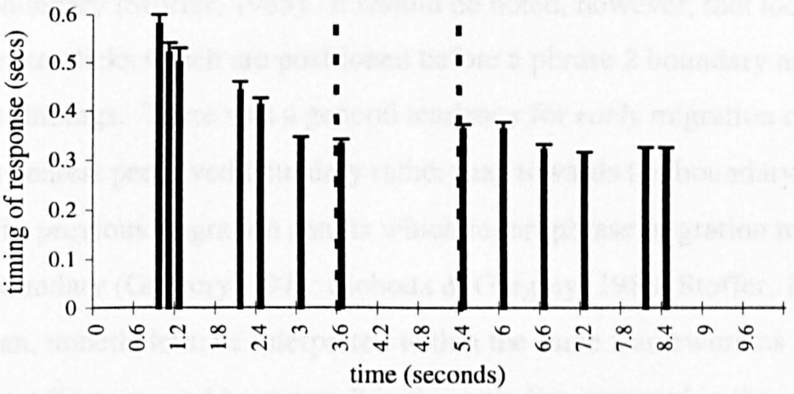


Figure 5c Click detection responses in phrase 1 -2 of melody S1.

Discussion

The click localisation technique was extended to two different melody excerpts using a wider range of click positions with a greater number of participants. The findings replicate those of the pilot study with late migration of clicks before boundary 1 and early migration of clicks after boundary 1 for both melodies. A comparison of click localisation performance of unfamiliar and familiar listeners revealed that there were no significant effects of familiarity on localisation responses in melodies H1 and S1. Localisation effects for clicks positioned before and after phrase boundaries were similar for unfamiliar and familiar listeners. This result suggests that listeners were picking up on the phrase structure of the melodies very quickly. The boundaries of these melodies are extremely salient, as was demonstrated in the Segmentation Experiment, above. Long notes in the melodies were salient boundary markers and phrases 1 - 2 of melodies S1 and H1 were grouped into two clear segments. Listeners may be able to detect phrase boundaries immediately without familiarisation of the melody.

The present localisation results are consistent with the theory that clicks which are placed prior to a phrase boundary tend to be migrated in the direction of that boundary and clicks which are placed after a boundary are migrated backwards in the direction of that boundary. This ties in well with previous results which have suggested that listeners try to maintain distinct perceptual units when processing incoming auditory information (Fodor & Bever, 1965; Swain, 1986). It is suggested that when listeners encounter a click in a piece of music they perceptually migrate it towards the nearest phrase boundary (Stoffer, 1985). It should be noted, however, that localisation responses to clicks which are positioned before a phrase 2 boundary are not in line with previous findings. There was a general tendency for *early* migration of clicks *away* from the nearest perceived boundary rather than towards the boundary. This is in contrast to previous migration results which report phrase migration towards the nearest phrase boundary (Gregory, 1978; Sloboda & Gregory, 1980; Stoffer, 1985). These results can, nonetheless, be interpreted within the same framework as previous studies. Migration effects around boundary 2 in the melodies occurred in the opposite direction than was expected. However, if we take into account the magnitude of click migration, the results may lend support to idea that discrete information units are perceived when listeners process auditory information. Listeners could have maintained perceptually discrete units by migrating the clicks backwards away from boundary 2, towards boundary 1. The finding that migration was more extreme the further from phrase 1 boundary and nearer phrase 2 boundary supports this suggestion. The unpredicted migration effects away from boundary 2 in both melodies may also be accounted for by considering memory factors. There is an increased memory load for listeners when required to respond to clicks which are positioned in the farthest temporal positions (i.e. those clicks nearest boundary 2 in both melodies). If the clicks occur early in the melody, the listener is able to stop actively processing the melody and can rehearse where the click occurs. This rehearsal is more difficult for clicks at the end of the melody as the rehearsal unit is too large for the listeners limited memory capacity (Atkinson & Shiffrin, 1968; Baddeley 1990). Listeners generally perceive groups of auditory information of 2 - 5 seconds duration (Fraisse, 1982). Clicks positioned nearer boundary 2 in melody S1 occur between 6 - 9 seconds of the excerpt. This is outwith the listeners immediate memory capacity and may explain the increasingly inaccurate localisation responses to clicks which occur farther into the melody.

A few points should be noted about the present results. Attention and memory factors may not be the only processes to account for the unexpected migration effects. It is possible that the unexpected migration effects may result from the way in which the stimuli excerpts were presented. In other words, although boundary 2 in each melody occurred later in time than boundary 1, boundary 2 also ended the experimental excerpts. It could have been the fact that boundary 2 ended the excerpt which contributed to the unexpected migration (away from boundary 2) which was observed in both melodies. This point is addressed in Condition 5, below, and will be discussed in turn. In addition to this, a broader examination of the data reveals a general trend for clicks to be perceived as occurring earlier than their actual position as the melody progresses, especially for melody H1. The relatively short notes of melody H1, and hence faster feel of the melody, may have caused the listener to anticipate the click occurrence as being much earlier than it in fact was. There is a serious problem of misinterpreting what may be a temporal effect as a perceived boundary effect at boundary 1 in melody H1. Boundary 1 of melody H1 was not utilised in any other experimental conditions which used the present click localisation task.

A comparison of localisation of the click *on* boundary 1 with clicks *before* and *after* boundary 1 yielded mixed results for melody S1. For phrase 1, responses to clicks *before* boundary 1 were not significantly different to the click positioned *on* the boundary. Similarly, there was no difference between localisation of clicks *before* boundary 2 and localisation of the click positioned *on* boundary 2. Responses to clicks *after* boundary 1 were significantly different to the click occurring *on* boundary 1. These results reveal that clicks which are positioned *after* boundary 1 were less accurately localised in comparison to the click occurring *on* the boundary. The results also show that localisation of the click positioned on boundaries was not more accurate than clicks positioned before the boundaries. In other words listeners respond to clicks positioned immediately before each boundary in the same way as they respond to the click on the boundary. The similarity of localisation response to clicks before and the click on the boundary suggests that listeners are anticipating the phrase boundary before its actual occurrence. This is an interesting finding, although a wider range of stimuli are needed to examine whether this effect is specific to these melodies.

Click *detection* results in melodies S1 and H1 highlight the effect of click position on detection time. Detection of the clicks in the first few seconds of each melody was slower than detection of clicks positioned later in the melodies. This may reflect that listeners are focused primarily on the melody rather than the click, as they were instructed to, especially for the first few seconds. The results may suggest that, as the melody progresses without click occurrence, listeners are aware of the increasing probability of click occurrence and rather than focus on the musical structure they anticipate the click. This could be problematic because one of the basic requirements of the click localisation task is that listeners attend to the melody primarily rather than the click sound (Stoffer, 1985). In order to ascertain whether listeners were focusing on the musical structure as directed, condition 3 examined click detection alone in melodies S1 and H1 without the subsequent localisation task. This was to enable a direct empirical comparison of click detection times between listeners who respond as soon as they detect the click (new Condition 3) and listeners whose primary task was to focus on the musical structure before detecting the click (the present Condition 2). The basic hypothesis is that there will be faster click detection in the *click detection only* group in comparison to the *click detection with subsequent localisation* group. This should be the case if listeners in Condition 2 were attending to the musical structure as directed rather than to the click alone.

Condition 3 Where *is* attention focused?

Method

Subjects The subjects were 11 psychology undergraduates, 6 females, 5 males, aged between 17 and 38 years. Participation in the experiment was in part fulfilment of a psychology course requirement.

Stimuli/Apparatus The stimuli were identical to that used in Condition 2, above.

Procedure The listener was required to perform a click detection task on hearing the click and was asked to subsequently listen to the repeated copy (without the click) which followed. Unlike Condition 2, above, no click localisation task was required. Practice trials were run using excerpts from two similar Gaelic melodies to enable the listener to become accustomed to the task. Experimental trials lasted an average of 40 minutes.

Click detection results

Click detection times were calculated in the same way as conditions 1 and 2, above. Mean detection times are displayed in Figures 6 and 7, below (and Tables 1 & 2, Appendix C, page 91). Performance on the click detection task without subsequent localisation task (referred to as Task B from now on) was compared with performance on the previous click detection task in Condition 2 (referred to as Task A from now on). The main difference between Task A and Task B was that the listener was required to attend to the click sound primarily in Task B rather than attend to the musical structure as in Task A.

Repeated-measures ANOVA's revealed a significant difference between click detection times for the two different tasks. Click position was the within-subjects factor and type of task was the between-subjects factor. For melody H1, there was a significant between-subjects effect, $F(1, 42) = 5.75$ at the $p < 0.05$ level. Clicks were detected an average of 0.07 seconds faster in Task B than in Task A. There was a significant main effect of position, $F(14, 588) = 7.22$, $p < 0.001$ and no interaction, $p > 0.90$.

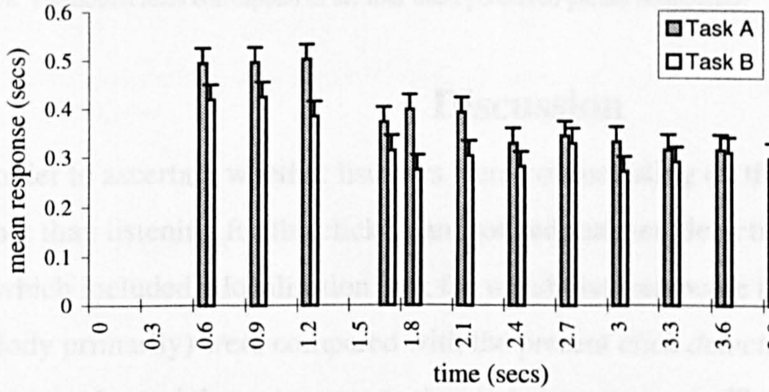


Figure 6 Mean click detection times for phrase 1 of melody H1 as a function of type of task.
The x-axis refers to the time of note occurrence in the melody. Each performance bar corresponds to a click placed in the melody. The y-axis shows the average time taken for listeners to detect the click. The two different tasks, Task A and Task B are described above.

This result was replicated in melody S1. There was a significant between-subjects effect with $F(1, 42) = 5.62$, $p < 0.05$. Clicks were detected an average of 0.05 seconds faster in Task B than in Task A. There was a significant effect of click position $F(28, 1176) = 8.58$, $p < 0.001$ and no interaction with $p > 0.90$. As can be seen from Figure 7, below, the position effect is due largely to the trend of faster detection of clicks which are positioned after note 5 in the melody, or which occur after 3 seconds into the melody.

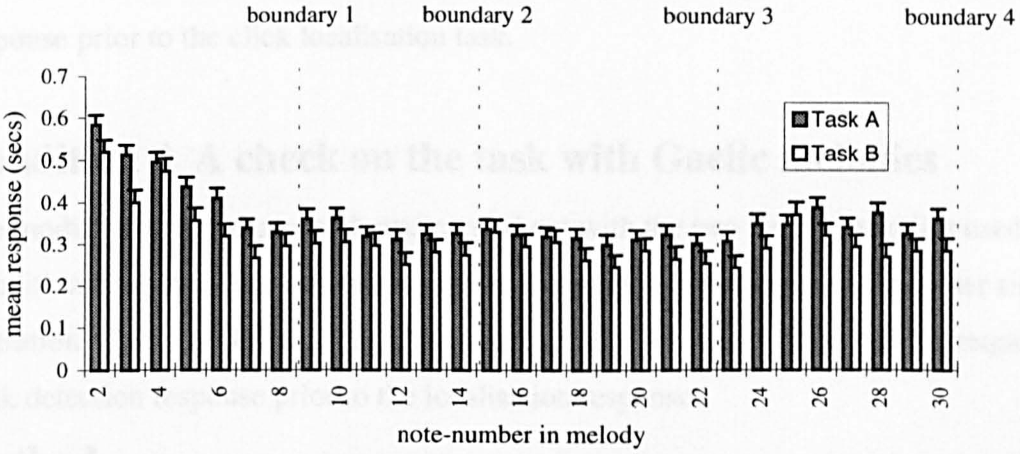


Figure 7 Mean click detection times for melody S1 as a function of type of task. The x-axis displays the note onset on which the click occurred, i.e. number 2 corresponds to note 2 in the melody and so on. The y-axis shows the average time taken for listeners to detect the click. The two different tasks are described above. The dotted lines correspond to the four main perceived phrase boundaries.

Discussion

In order to ascertain whether listeners were concentrating on the music as requested rather than listening for the click in an isolated manner, detection results from Condition 2 (which included a localisation task for which listeners were instructed to focus on the melody primarily) were compared with the present *click detection only* results. Analyses showed that responses to the *click detection only* (Task B) condition were significantly faster than responses to the *click detection when focusing on musical structure* (Task A) condition. Click detection was an average of 0.05 seconds faster in melody S1 and 0.07 seconds faster in melody H1 in Task B (click detection only) compared with Task A (click detection and localisation task). The different responses on these two detection tasks suggests that listeners in Task A were concentrating on the melody rather than listening for the click in an isolated manner. As a result of paying attention to structural features in the music, click detection times took longer. This implies that listeners were attending to the music as instructed in Condition 2 (Task A) above.

One remaining concern with the present technique is that the additional click detection task may influence performance on subsequent localisation performance. Condition 4 examined this further using a modified task which did not require a click detection response prior to the click localisation task.

Condition 4 A check on the task with Gaelic melodies

The modified experimental task was carried out with the two Gaelic melodies used in Condition 1, above. The aim of the present experiment was to examine whether similar migration effects are found using a modified localisation task which does not require a click detection response prior to the localisation response.

Method

Subjects The subjects were 20 psychology undergraduates, 14 females, 6 males, aged between 17 and 42 years. Participation in the experiment was in part fulfilment of a psychology course requirement.

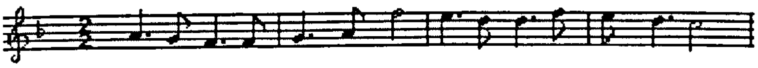
Stimuli/Apparatus The stimuli were phrases 1 - 2 of melodies S2 and H2, shown below in Figure C.

Figure C

Melody S2 phrase 1 - 2

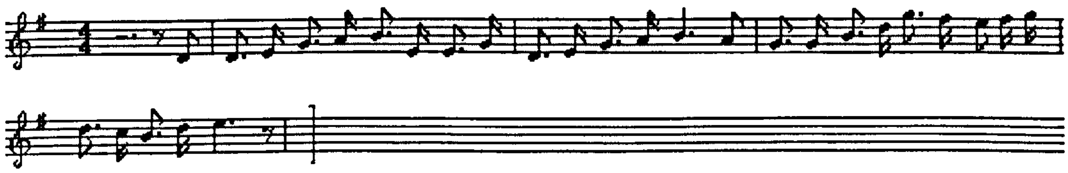
Perceived boundary number: 1 2

Click position 1 2 3 4 5 6 7 8 9 10 11 12



Melody H2, phrase 1 - 2

Perceived boundary number: 1

Click position 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21

22 23 24 25 26

boundary number: 2

Procedure Listeners were required to listen to the excerpt containing the click and then perform a click localisation task during the repeated copy without the click which followed. Unlike Conditions 1 and 2, there was no click detection task in the present experiment. Half the listeners performed with melody S2 and half with melody H2. Practice trials were run using excerpts from two similar Gaelic melodies to enable the listener to become accustomed to the task. Experimental trials lasted an average of 30 minutes.

Click localisation results

Mean localisation responses are shown in Figures 8a and 9a, below (and in Tables 1 & 2, Appendix D, page 92). A repeated measures ANOVA was performed on the data with click position as the within-subjects factor. Each melody was analysed separately.

Melody H2

The ANOVA revealed a significant effect of click position on localisation response, F(25, 200) = 2.32, $p < 0.001$. Mean responses are shown in Figure 8a, below.

Responses were then grouped into localisation of clicks positioned on notes within 2 crotchet beats before boundary 1 and within 2 crotchet beats after boundary 1 for further analysis. Means are displayed in Figure 8b, below.

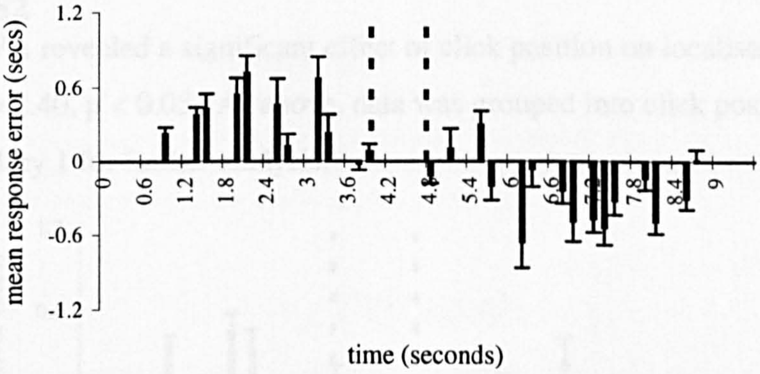


Figure 8a Click localisation response to phrases 1 - 2 of melody H2
The x-axis represents time in seconds, each bar corresponds to a click positioned on a note in the melody. The dotted lines mark the onset and offset of boundary 1. The y-axis shows the mean localisation response in seconds. A response at zero indicates a correct response, a positive response time represents a late migration and a negative response time represents an early migration.

Again, the main interest lay in examining click migration around perceived boundary 1. There was no significant difference between localisation of clicks *before* and *after* boundary 1, $p > 0.10$. The pattern for more accurate localisation of clicks directly before boundary 1 (reported in Conditions 1 and 2, above) is replicated in the present experiment (Figure 8b, below). There was no difference between localisation of clicks *before* and the click *on* boundary 1, $p > 0.10$. Similarly, there was no difference between localisation of the click positioned *on* and clicks positioned *after* boundary 1, p > 0.70. There was no difference between localisation of clicks *before* and the click *on* boundary 2, p > 0.20.

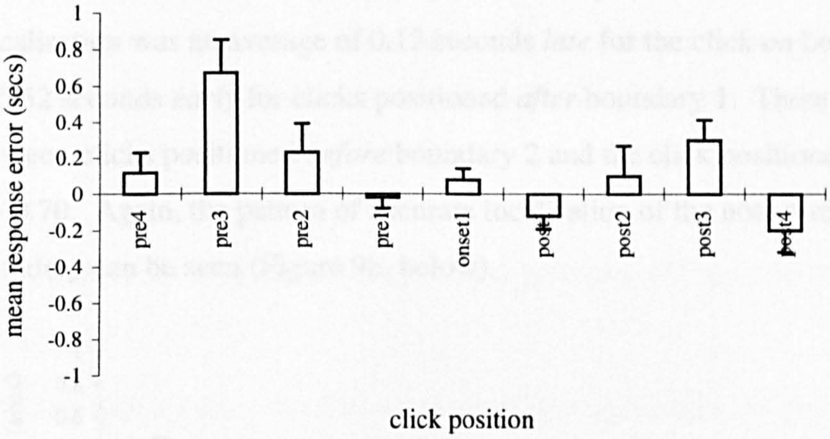


Figure 8b Magnified migration effects around boundary 1 in melody H2
The y-axis shows localisation response and the x-axis shows the position of click on notes before, on and after boundary 1. Pre4 refers to a click positioned on the fourth note before the boundary and post 4 refers to a click positioned on the fourth note after boundary 1 and so on.

Melody S2

The ANOVA revealed a significant effect of click position on localisation response, F(11, 99) = 2.40, $p < 0.05$. As above, data was grouped into click positions before and after boundary 1 for further analysis.

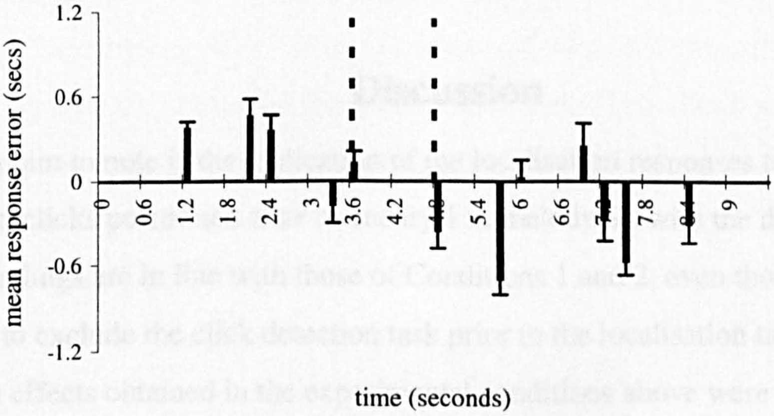


Figure 9a Click localisation responses to phrase 1- 2 of melody S2
As detailed above in Figure 8a, the x-axis represents time in seconds and the y-axis shows the mean localisation response in seconds. The dotted lines correspond to the onset and offset of boundary 1.

A repeated measures ANOVA performed on the grouped before and after boundary 1 data revealed a significant migration effect, F(1, 9) = 7.02, $p < 0.05$. Localisation of clicks *before* boundary 1 were migrated an average of 0.15 seconds *later* than actual occurrence whereas clicks positioned *after* boundary 1 were migrated an average of 0.52 seconds *earlier* than actual occurrence. Localisation of the click positioned *on* boundary 1 was no different to localisation of clicks positioned *before* boundary 1, $p > 0.80$. Localisation of clicks positioned *after* boundary 1 was significantly different to localisation of the click positioned *on* boundary 1, $F(1, 9) = 5.93$, $p < 0.05$. Localisation was an average of 0.12 seconds *late* for the click *on* boundary 1 in contrast to 0.52 seconds *early* for clicks positioned *after* boundary 1. There was no difference between clicks positioned *before* boundary 2 and the click positioned *on* boundary 2, $p > 0.70$. Again, the pattern of accurate localisation of the note directly before the boundary can be seen (Figure 9b, below).

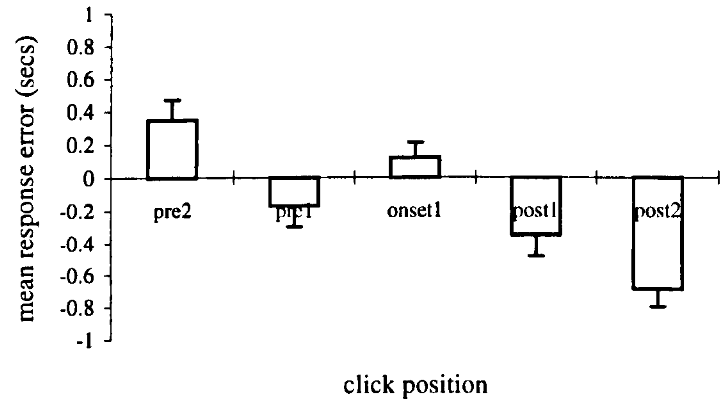


Figure 9b Magnified migration effects around boundary 1 in melody S2
As detailed in Figure 8b, above, the x-axis displays the click position on notes before, on and after boundary 1. The y-axis shows the mean localisation response.

Discussion

The first point to note is the replication of the localisation responses to clicks positioned before and clicks positioned after boundary 1 in melody S2 with the different task. The present findings are in line with those of Conditions 1 and 2, even though the task was modified to exclude the click detection task prior to the localisation task. Thus the migration effects obtained in the experimental conditions above were not the result of the particular experimental task, nor were they due to the particular melodies used. Click detection prior to the click localisation task is not necessary to obtain a migration response. In addition to this, the trend of more accurate localisation for clicks which are positioned on the penultimate note in the first perceived phrase is again found here in both melodies. Also in line with the findings from conditions 1 and 2 is the general tendency for localisation responses to clicks before boundary 2 to migrate in the opposite direction of boundary 2. As mentioned above, pp 61 - 62, it is necessary to investigate whether the actual excerpts cause these unexpected migration effects. In other words, do the unpredicted localisation responses *away* from boundary 2 occur because boundary 2 ended the excerpt. To examine whether this was the case, Condition 5 was carried out to compare migration effects in melodies S2 and H2 for clicks positioned on notes before boundary 2 *ending the excerpt* and clicks positioned on the same notes before boundary 2 but *not* ending the excerpt.

Condition 5 Are unexpected migration effects stimulus specific?

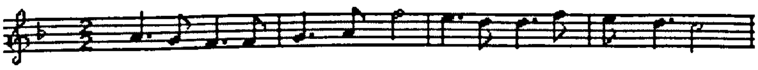
Method

Subjects The subjects were 20 psychology undergraduates, 13 females, 7 males, aged between 19 and 31 years. Participation in the experiment was in part fulfilment of a psychology course requirement.

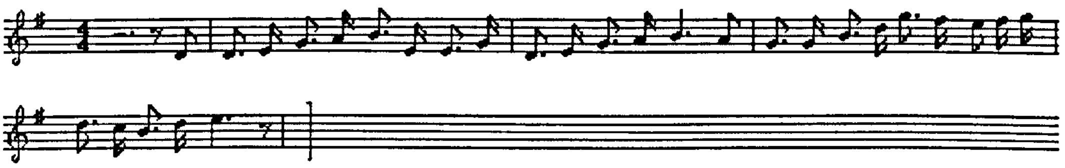
Stimuli/Apparatus The stimuli were phrases 1 - 2 and phrases 2 -3 of melodies H2 and S2, shown below (Figure D). Clicks were positioned on the notes numbered 1 - 4 in melody S2 and notes numbered 1 - 7 in melody H2. Phrase 1 - 2 excerpts were the "ending the excerpt" condition in which the click was positioned on a note which occurred near the end or on the end of the excerpt. Clicks were also positioned on the same notes, this time as part of the phrase 2 - 3 excerpt. This created the "not ending the excerpt" condition in which the click was positioned on a note which did not end the excerpt. The overall duration of the "ending the excerpt" condition was identical to the "not ending the excerpt" conditions (9.6 seconds for "ending" and 9.6 seconds for "not ending" for both melodies). Click position, although on exactly the same notes, changed in relation to position in the overall excerpt. Localisation of clicks in the two conditions ("ending the excerpt" and "not ending the excerpt") were compared.

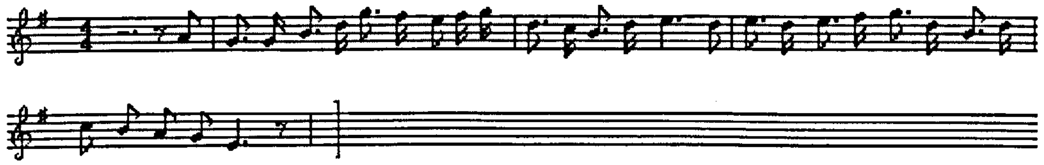
Figure D Melody excerpts used to examine the effect of excerpt ending on click localisation.

Melody S2 phrases 1 - 2 **"Ending the excerpt"**
Click position: 1 2 3 4



Melody S2 phrases 2 - 3 **"Not ending the excerpt"**
Click position: 1 2 3 4

Melody H2 phrases 1 - 2 **"Ending the excerpt"**
Click position: 1 2

3 4 5 6 7

Melody H2 phrases 2 - 3 **"Not ending the excerpt"**
Click position 12 3 45 6 7

Procedure Listeners performed the same task as in Condition 4, above. Half the listeners performed with melody S2 and half with melody H2. Practice trials were run using excerpts from two similar Gaelic melodies to enable the listener to become accustomed to the task. Experimental trials lasted an average of 15 minutes.

Click localisation results

Melodies S2 and H2

Localisation of clicks positioned on notes which ended the excerpt were compared with localisation of clicks on these same notes, this time not ending the musical excerpt. A repeated-measures ANOVA was performed on the two sets of grouped responses with melody type as the between-subjects factor and click position as the within-subjects factor (this was either "ending the excerpt" or the same notes "not ending the excerpt"). Means are displayed in Figure 10, below (and Table 1, Appendix E, page 93). The results revealed a significant effect of click position in the excerpt, $F(1, 18) = 12.59$, $p < 0.005$, no effect of melody $p > 0.70$ and no melody type-click position interaction, $p > 0.20$. Clicks positioned on notes which ended the excerpt were migrated backwards an average of 0.27 seconds, i.e. they were perceived as occurring *earlier* than their correct position. The clicks positioned on the same group of notes, this time not ending the excerpt, were migrated an average of 0.18 seconds towards the phrase boundary, in other words they were perceived as occurring *later* than their actual position.

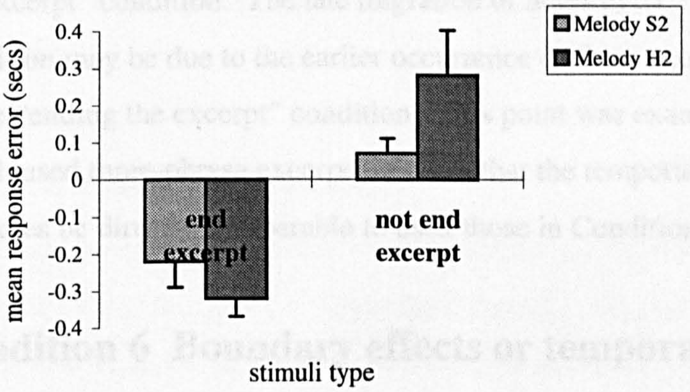


Figure 10 Localisation response as a function of note position in overall excerpt. The y-axis shows mean response error and the x-axis displays the two stimulus groups. "end excerpt" refers to the notes which ended the excerpt and "not end excerpt" refers to the same notes which did not end the excerpt.

Discussion

The main finding was that localisation was significantly different for clicks positioned on notes which ended an excerpt compared with clicks positioned on the same notes but which did not end the excerpt. If the clicks were close to the end of the musical excerpt then they were migrated *away* from the nearest major phrase boundary. This was the result found in Conditions 2 & 4, above, in which clicks positioned before boundary 2 (which ended the excerpt) were migrated away from boundary 2. Clicks which were positioned on notes *not ending* the excerpt were migrated *towards* the nearest phrase boundary. This finding has implications for the unexpected migration effects away from the nearest phrase boundary which occurred in conditions 1, 2 and 4. It suggests that the backwards or early migration away from the second phrase boundaries could have resulted because those phrases ended the excerpt. The present experiment reveals that boundary 2 in melodies H2 and S2 attracts click migration in the expected direction when it does not end the excerpt. This, however, is not a complete account of the unexpected migration effects obtained in conditions 1, 2 and 4. As mentioned above, it is possible that temporal and attentional factors may affect click migration in melodies. By the time listeners perceive clicks near boundary 2 the amount of melody to remember has exceeded the listeners memory capacity. This may cause the increasingly erroneous early migration of the clicks which are positioned nearer boundary 2. The present experiment does not exclude the possibility of a temporal effect at boundary 2. Clicks in the "not ending the excerpt" condition occur much earlier than in the "ending the excerpt" condition. The late migration of notes in the "not ending the excerpt" condition may be due to the earlier occurrence of these notes in comparison to the notes in the "ending the excerpt" condition. This point was examined further in Condition 6 which used three-phrase excerpts in order that the temporal position of the clicks in the melodies be directly comparable to used those in Condition 4, above.

Condition 6 Boundary effects or temporal effects?

Method

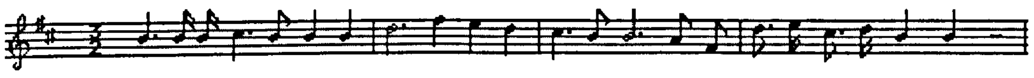
Subjects The subjects were 49 psychology undergraduates, 31 females, 18 males, aged between 18 and 38 years. Participation in the experiment was in part fulfilment of a psychology course requirement.

Stimuli/Apparatus The stimuli were phrases 1 - 3 of melodies S1, S2 and H2, shown below in Figure E. Apparatus was the same as in the experiments above.

Figure E Melodies used in the present condition

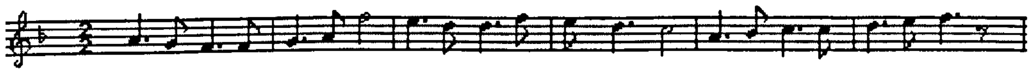
Melody S1

perceived boundary number: 1 2

Click position: 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21

Melody S2

perceived boundary number: 1 2

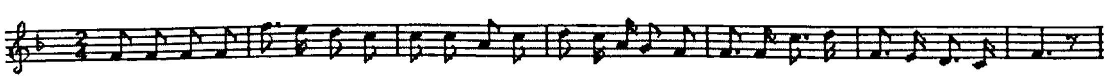
Click position: 1 2 3 4 5 6 7 8 9 10 11 12 13 14



Melody H1

perceived boundary number: 2

Click position: 1 2 3 4 5 6 7 8 9



Procedure Listeners performed the same task as in Condition 1, above. Fourteen of the listeners performed with melody S1, fifteen with melody S2 and twenty with melody H1. Practice trials were run using excerpts from two similar Gaelic melodies to enable the listener to become accustomed to the task. Experimental trials lasted between 20 and 30 minutes.

Click localisation results

Localisation responses to clicks in each melody are shown below in Figures 11a, 12a & 13a (and Tables 1a, 2a & 3a, Appendix F, pp 93 - 94). Localisation responses to each melody are analysed separately. The data was grouped into responses to clicks positioned on notes two crotchet beats before and after boundaries 1 and 2 in each melody for further analysis. This resulted in the following four separate groups for melodies S1 and S2:

1. Group pre-phrase 1: the two notes before boundary 1
2. Group post-phrase 1: the two notes after boundary 1
3. Group pre-phrase 2: the two notes before boundary 2
4. Group post-phrase 2: the two notes after boundary 2 for S2 and the four notes after boundary 2 for S1

and the following two groups for phrase 2 of melody H1:

1. Group pre-phrase 2: the four notes before boundary 2
2. Group post-phrase 2: the four notes after boundary 2

Melody S2

A repeated-measures ANOVA was performed on the before and after boundary data. This revealed a significant effect of click position on localisation response for boundary 1, $F(1, 14) = 8.72$, $p < 0.01$. Localisation of clicks positioned *before* boundary 1 was an average of 0.33 seconds *later* than actual click occurrence. Localisation of clicks positioned directly *after* boundary 1 was an average of 0.20 seconds *earlier* than actual position. Similar results were found for boundary 2, $F(1, 14) = 11.94$, $p < 0.005$. Localisation of clicks *before* boundary 2 were migrated an average of 0.04 seconds *later* than actual click position whilst responses to clicks positioned *after* boundary 2 were migrated an average of 0.67 seconds *earlier* than correct position (means are displayed in Figures 11a & 11b, below).

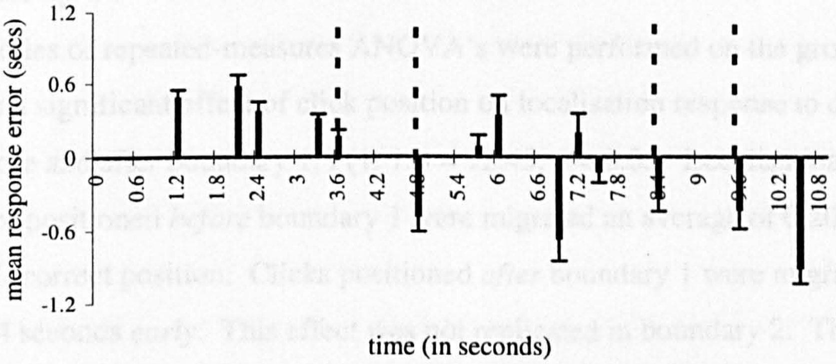


Figure 11a Localisation responses to phrases 1 - 3 in melody S2.
The y-axis shows the mean response error. A response of zero is a correct localisation, a positive response is an early localisation and a negative response is a late localisation. Each bar on the x-axis corresponds to a note on which a click was positioned. The dotted lines mark the onset and offset of perceived boundaries 1 and 2.

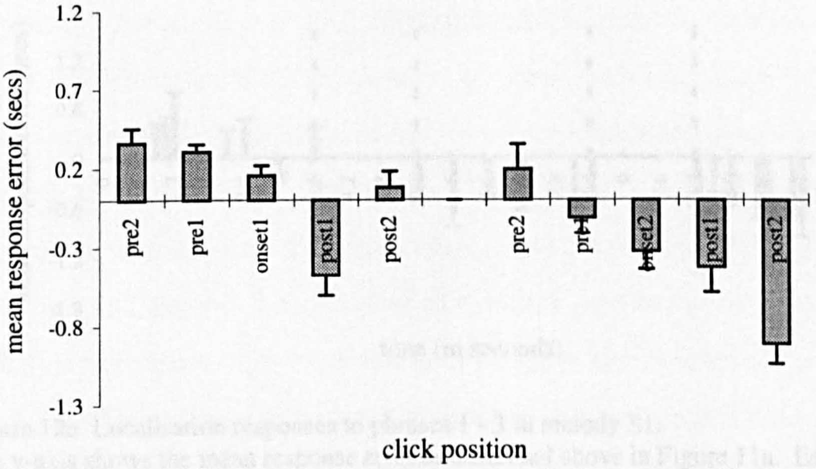


Figure 11b Magnification of localisation response to clicks positioned on notes *before*, *on* and *after* boundaries 1 and 2 in melody S2.
The y-axis shows the mean response error as described above in Figure 11a. The x-axis displays the notes on which the clicks were positioned, e.g. pre1 refers to the note before the boundary, onset 1 to the boundary 1 note and post2 to the second note after the boundary.

Comparisons were made between localisation of the click *on* boundaries 1 and 2 and localisation of the clicks on notes immediately *before* and immediately *after* those boundaries, using the groups identified above. Localisation of the click positioned *on* boundary 1 was not significantly different to localisation of clicks positioned on notes *before* boundary 1, $p > 0.20$ or localisation of clicks positioned on notes *after* boundary 1, $p > 0.06$. Similar results were found for boundary 2. Localisation of the click positioned on boundary 2 was not significantly different to localisation of clicks positioned on notes *before* boundary 2, $p > 0.20$ and *after* boundary 2, $p > 0.30$.

Melody S1

A series of repeated-measures ANOVA's were performed on the grouped data. There was a significant effect of click position on localisation response to clicks positioned before and after boundary 1, $F(1, 13) = 12.45$, $p < 0.05$. Localisation of clicks which were positioned *before* boundary 1 were migrated an average of 0.20 seconds *later* than their correct position. Clicks positioned *after* boundary 1 were migrated an average of 0.54 seconds *early*. This effect was not replicated in boundary 2. There was no difference between localisation of clicks positioned before and clicks positioned after boundary 2, $p > 0.90$. Localisation of clicks *before* and *after* boundary 2 migrated and average of 0.58 and 0.57 seconds *early*, respectively. Means are displayed in Figure 12a & 12b, below.

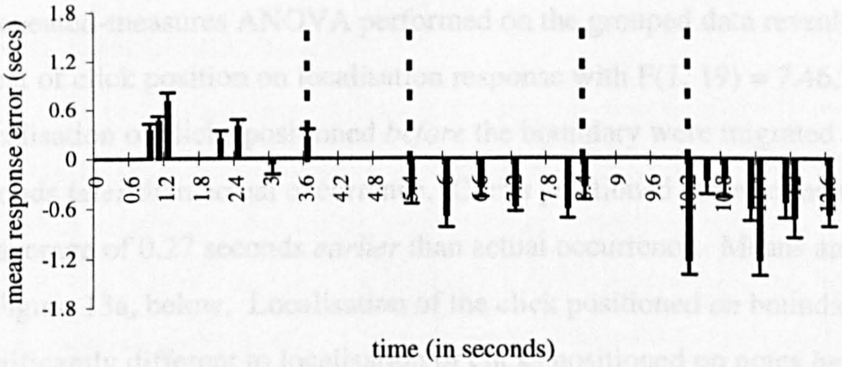


Figure 12a Localisation responses to phrases 1 - 3 in melody S1.
The y-axis shows the mean response error as described above in Figure 11a. Each bar on the x-axis corresponds to a note on which a click was positioned. The dotted lines mark the onset and offset of perceived boundaries 1 and 2.

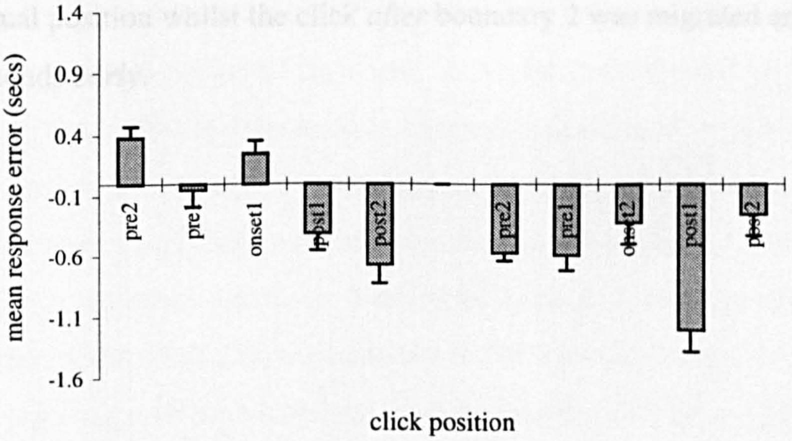


Figure 12b Magnification of localisation response to clicks positioned on notes *before*, *on* and *after* boundaries 1 and 2 in melody S1.

Localisation of the click positioned *on* boundary 1 was not significantly different to localisation of clicks positioned on notes *before* boundary 1, $p > 0.70$. The result was the same for boundary 2, with no difference between localisation of clicks positioned *before* boundary 2 and localisation of the click positioned *on* boundary 2. There was a significant difference between localisation of the click *on* boundary 1 and localisation of clicks positioned *after* boundary 1, $F(1, 13) = 12.45$, $p < 0.005$. The click *on* boundary 1 was migrated an average of 0.25 seconds *later* than actual position and the clicks positioned *after* boundary 1 were migrated an average of 0.54 seconds *early*. This significant difference between the click *on* and clicks *after* the boundary was not replicated for boundary 2, with $p > 0.60$.

Melody H1

A repeated-measures ANOVA performed on the grouped data revealed a significant effect of click position on localisation response with $F(1, 19) = 7.46$, $p < 0.05$. Localisation of clicks positioned *before* the boundary were migrated an average of 0.04 seconds *later* than actual occurrence. Clicks positioned *after* boundary 2 were migrated an average of 0.27 seconds *earlier* than actual occurrence. Means are displayed below in Figure 13a, below. Localisation of the click positioned *on* boundary 2 was not significantly different to localisation of clicks positioned on notes *before* boundary 2, $p > 0.20$. There was a significant difference between localisation of the click on boundary 2 and localisation of clicks positioned *after* boundary 2, $F(1, 13) = 20.76$, $p < 0.001$. The click *on* boundary 2 was migrated an average of 0.17 seconds *later* than actual position whilst the click *after* boundary 2 was migrated an average of 0.27 seconds *early*.

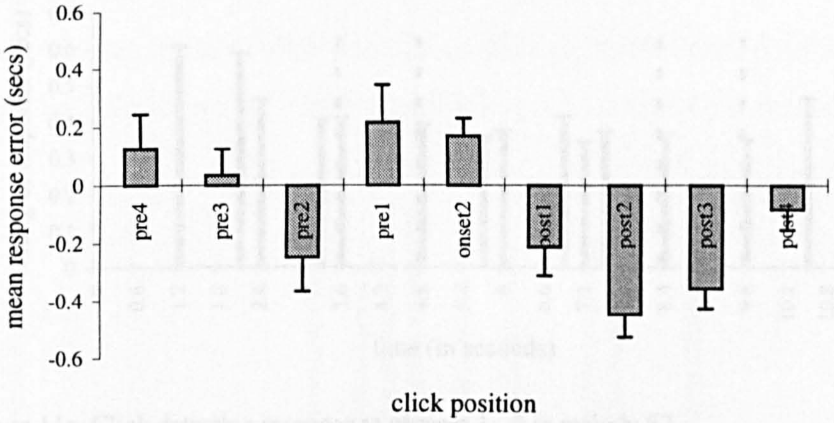


Figure 13a Localisation responses to clicks positioned around boundary 2 in melody H1.
As described above, in Figure 11b, the y-axis displays mean localisation response and the x-axis shows the note on which the click was positioned.

Click detection results

Click detection times were calculated in the same way as Conditions 1, 2 and 3, above. A series of repeated measures ANOVA's were performed on the click detection data with click position as the within subject factor. Mean detection times are displayed in Figures 11c, 12c & 13c (and Tables 1b, 2b, & 3b, Appendix F, pp 93 - 95) for melodies H1, S1 and S2, respectively. Click detection responses to each melody are discussed in turn.

Melody S2

The ANOVA revealed a significant effect of click position on detection time, F(1, 14) = 3.89, $p < 0.001$. Figure 11c, below, shows the general trend for faster detection of clicks the farther into the melody the click is positioned, especially for the first 3 seconds of the melody. Detection of clicks which were positioned *on* perceived boundaries 1 and 2 and *immediately after* those boundaries were then compared with detection of clicks *inbetween* boundaries 1 and 2. It was hypothesised that information in the melody would be higher at the end of a phrase and the beginning of the next phrase compared with information in the middle of a phrase. There were, however, no significant differences between these two groups, $p > 0.20$. Thus detection times were similar for clicks occurring on the two perceived boundary notes and clicks positioned on notes inbetween these two boundaries.

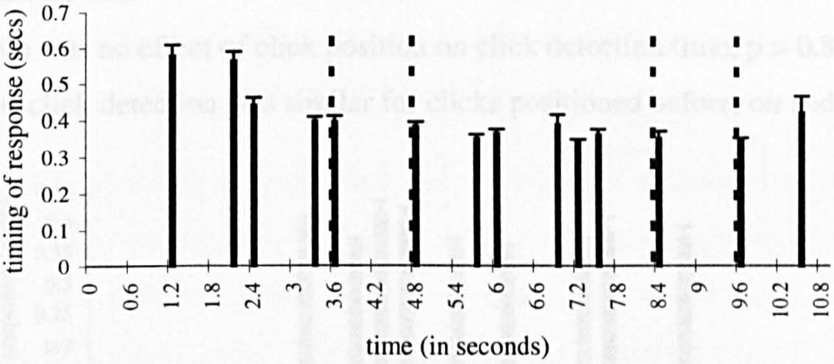


Figure 11c Click detection response to phrases 1 - 3 in melody S2.
The y-axis shows the detection time in seconds and bars on the x-axis correspond to notes on which clicks were positioned. The dotted lines correspond to the onset and offset of perceived boundaries 1 and 2.

Melody S1

The ANOVA revealed a significant effect of click position on detection time, F(1, 13) = 8.18, $p < 0.001$. Figure 12c, below, shows the general trend of faster click detection for clicks farther into the melody for the first 2 seconds. Detection of clicks which were positioned *on* perceived boundaries 1 and 2 and *immediately after* those boundaries were then compared with detection of clicks *inbetween* boundaries 1 and 2. There was a significant difference between effect of click position, F(1, 13) =15.63, $p < 0.001$. Detection of clicks positioned on notes in between boundaries 1 and 2 was significantly faster by an average of 0.08 seconds than detection of clicks on boundaries 1 and 2.

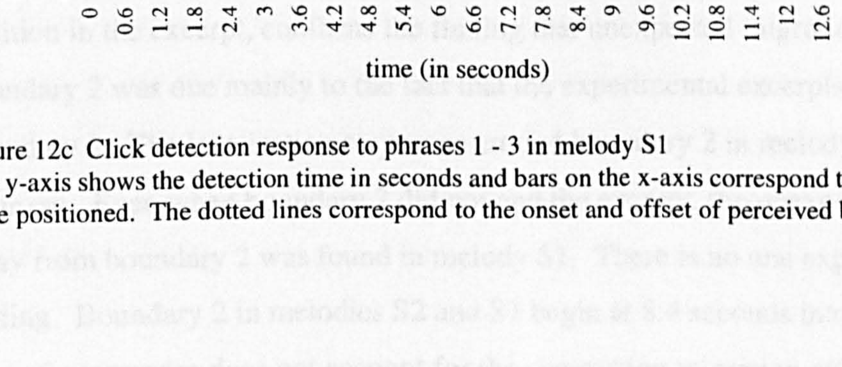


Figure 12c Click detection response to phrases 1 - 3 in melody S1
The y-axis shows the detection time in seconds and bars on the x-axis correspond to notes on which clicks were positioned. The dotted lines correspond to the onset and offset of perceived boundaries 1 and 2.

Melody H1

There was no effect of click position on click detection time, $p > 0.80$ for melody H1. Thus click detection was similar for clicks positioned before, on and after boundary 2.

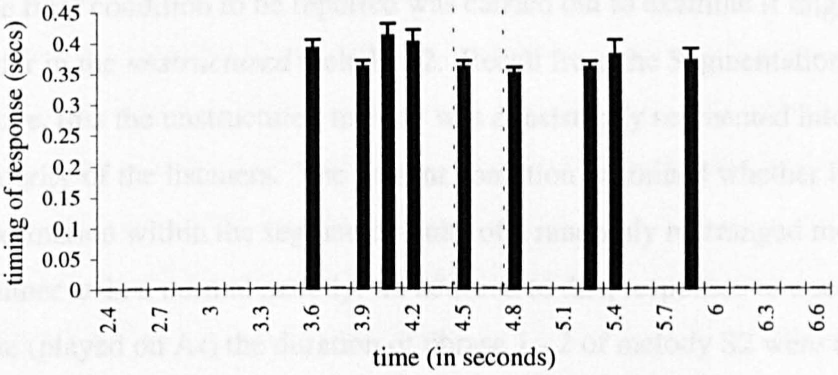


Figure 13c Click detection response around boundary 2 in melody H1
The y-axis shows the detection time in seconds and bars on the x-axis correspond to notes on which clicks were positioned. The dotted lines correspond to the onset and offset of perceived boundary 2.

Discussion

The present experiment replicated the migration effects around boundary 1 in melodies S1 and S2 which were reported above in Conditions 2 and 4, respectively. Analyses also revealed that clicks before and after boundary 2 in melodies S2 and H1 are migrated in the expected direction, i.e. towards the main perceived boundary 2 in each melody. These findings are in contrast to results obtained in Condition 2 in which clicks positioned on the same notes in the same temporal positions were migrated away from boundary 2. The excerpts used in Condition 2 ended on boundary 2. Condition 4 suggested that the unexpected migration away from boundary 2 was due mainly to the fact that excerpts ended on boundary 2, although an effect of temporal position could not be ruled out. The present experiment, which takes into account overall temporal position in the excerpt, confirms the finding that unexpected migration away from boundary 2 was due mainly to the fact that the experimental excerpts ended on boundary 2. The localisation responses around boundary 2 in melody S1 are not as clear cut. Even when boundary 2 did not end the excerpt, the unexpected migration away from boundary 2 was found in melody S1. There is no one explanation for this finding. Boundary 2 in melodies S2 and S1 begin at 8.4 seconds into the melody so the time of occurrence does not account for the contrasting migration effects. The overall duration of the melody S1 (12.6 seconds) excerpt is however longer than melody S2 (10.5 seconds). The additional memory load of the longer excerpt may have contributed to the unexpected migration effects at boundary 2.

The final condition to be reported was carried out to examine if migration effects also occur in the *unstructured* melody S2. Recall from the Segmentation Experiment, above, that the unstructured melody was consistently segmented into two groups by a majority of the listeners. The present condition examined whether listeners process information within the segmented units of a randomly rearranged melody in the same manner as in a normal melody. In addition to this, responses to a second stimulus, one note (played on A4) the duration of phrase 1 - 2 of melody S2 were also examined. This note lasted 9.6 seconds (as did phrase 1- 2 of melody S2) and contained a click positioned in the same temporal position as the clicks positioned in melody S2, Condition 4, above. The same 12 click positions used in Condition 4 were used 12 times with the long note. Responses to the click in the note were compared with responses to the corresponding click in melody S2. This condition was carried out to examine whether localisation responses were temporally based or related to aspects of the musical structure. If responses were different between the two conditions then this suggests that it was the musical structure which influenced the migration responses in Condition 4.

Condition 7 Migration effects in unstructured melody S2 and one tone lasting the duration of melody S2

This experiment examined whether migration effects occur in the expected direction at perceived boundaries in an unstructured melody. It also examined whether responses to the click positioned in the long note (which lasted the duration of phrase 1 - 2 of a melody) are similar to responses to the corresponding click positioned in the real melody. The results of localisation responses to clicks which were positioned within the long held note were compared with responses to melody S2 which were obtained in Condition 4, above.

Method

Subjects The subjects were 33 psychology undergraduates, 19 female, 14 male, aged between 18 and 42 years. Participation was in part fulfilment of a psychology course requirement.

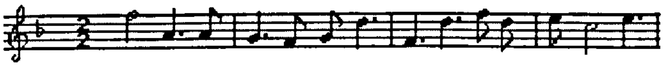
Stimuli/Apparatus The stimuli were excerpts of unstructured melody S2 which consisted of phrase 1-2 of melody S2 but this time the order of note occurrence was rearranged. All events in the melody were preserved (the pitch and duration of notes was exactly the same) but the order was randomised. The other stimuli were excerpts of one tone, A4, which lasted the duration of phrase 1-2 of melody S2. The tone contained a click positioned in the same temporal position as a click in melody S2. There were twelve tone conditions. Each contained a click positioned in a position which corresponded to a position used in phrase 1- 2 of melody S2. The experimental stimuli are shown in Figure F, below.

Procedure Listeners performed the same task as Condition 4, above. Fifteen listeners performed with melody S2 unstructured and eighteen with the tone. Practice trials were run using excerpts from two similar Gaelic melodies to enable the listener to become accustomed to the task. Experimental trials lasted between 20 and 30 minutes.

Figure F

Unstructured melody S2, phrase 1 - 2

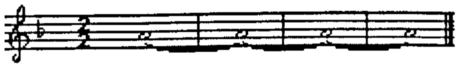
Click position 1 2 3 4 5 6 7 8 9 10 11 12 13



perceived boundary number: 1 2

Comparison tone the duration of phrase 1 - 2 of melody S2

Click positions 1 - 12 on the same temporal positions on which clicks occur in melody S2 (Figure C, Condition 4, above page 66)



Click localisation results

Melody S2 unstructured

Mean localisation responses are shown in Figures 14a and 14b, below (and Table 1, Appendix G, page 95). A repeated measures ANOVA revealed a significant effect of click position on localisation response, $F(12, 156) = 3.77$, $p < 0.001$. Localisation responses were grouped into those made to clicks positioned on notes within 2 crotchet beats before boundary1 and those made to clicks positioned within 2 crotchet beats after boundary 1 for further analysis.

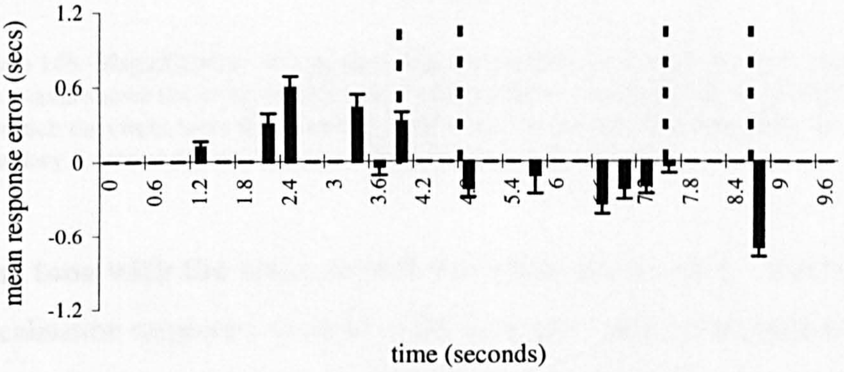


Figure 14a Migration responses to unstructured melody S2.
The x-axis represents time in seconds. Each response bar corresponds to a click positioned on a note in the unstructured melody. Dotted lines correspond to the onset and offset of the two perceived boundaries. Mean localisation times are shown on the y-axis.

A series of repeated measures ANOVA's revealed a significant difference between localisation of clicks positioned on notes before boundary 1 and clicks positioned on notes after boundary 1, $F(1, 14) = 5.32$, $p < 0.05$. Clicks positioned *before* boundary 1 were migrated an average of 0.33 seconds *later* than their actual position. Clicks positioned *after* boundary 1 were migrated an average of 0.19 seconds *earlier* than their actual position. Comparisons were made between localisation responses to the click positioned *on* boundary 1 and localisation of clicks positioned *before* and *after* boundary 1. There was no difference between localisation of clicks positioned before and the click positioned on boundary 1, $p > 0.90$. There was a significant difference between localisation of clicks positioned *after* boundary 1 and the click positioned *on* boundary 1, $F(1, 14) = 7.63$, $p < 0.05$. Clicks positioned after the boundary were migrated an average of 0.19 seconds *early* and those positioned on the boundary were migrated an average of 0.32 seconds *later* than actual position. There was no difference between localisation of clicks positioned *before* boundary 2 and localisation of the click positioned *on* boundary 2, p > 0.20.

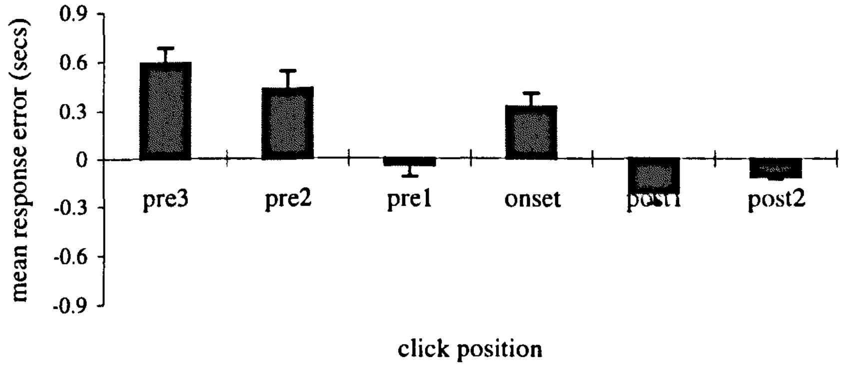


Figure 14b Magnification of migration responses around boundary 1 in unstructured melody S2. The y-axis shows the mean response error as described above in Figure 11a. The x-axis displays the notes on which the clicks were positioned e.g. pre3 refers to the third note before the boundary, onset to the boundary 1 note and post2 to the second note after the boundary.

One tone with the same overall duration as phrase 1 - 2 of melody S2

Localisation responses to clicks in the long note (which contained clicks in the same temporal positions as those in phrase 1 - 2 of melody S2) were compared with localisation responses to clicks in melody S2 (Condition 4, above). A repeated measures ANOVA was performed on the two groups of data with click position as the within subjects factor and participant group (melody S2 in Condition 4, or the long note) as the between subjects factor. There was a significant effect of participant group, F(1, 26) = 5.16, p < 0.05, a significant main effect of click position, F(11, 286) = 4.23, p < 0.001 and a significant group - click position interaction, F(11, 286) = 2.56, p < 0.005. Localisation of clicks which were positioned in melody S2 was significantly different to localisation of clicks which were temporally positioned in the identical places within a tone lasting the duration of melody S2. Mean localisation responses to the note are displayed in Figures 15a & 15b, below (and in Table 2, Appendix G, p. 95).

An additional analysis was run comparing responses to clicks in the long note positioned at the same time as notes before and after boundary 1 of melody S2 (because the group - click interaction was statistically significant). This revealed that there was no significant difference between localisation of clicks positioned in the long note *at the same time* as clicks positioned before and after boundary 1 of melody S2, p > 0.09.

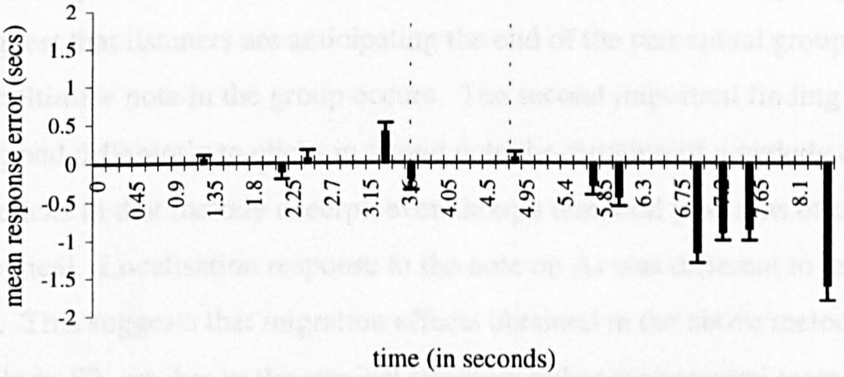


Figure 15a Migration effects in response to clicks within one sounded tone.
The x-axis shows time, in seconds, with each response bar corresponding to a click position in the sound burst. The y-axis represents the mean response error in seconds. Dotted lines correspond to the onset and offset of what would have been the boundary 1 marker in melody S2.

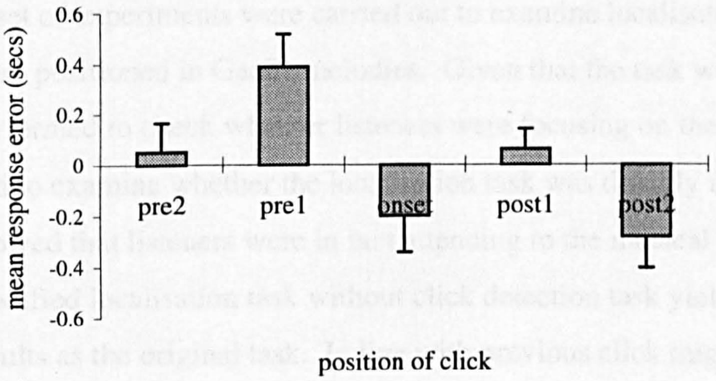


Figure 15b Magnification of migration effects in response to clicks within one sounded tone.
The x-axis corresponds to the notes before, on and after what would have been the boundary 1 offset in melody S2. The y-axis represents the mean response error in seconds.

Discussion

This experiment revealed two important findings. First of all, localisation responses to the unstructured melody were similar to those of normal melodies. There was a significant difference between localisation of clicks on notes before perceived boundary 1 and localisation of clicks on notes after perceived boundary 1. In addition to this, the results replicated the finding of more accurate localisation of clicks which were positioned on the penultimate note in the first perceived phrase. This suggests that listeners process information within segments in unstructured melodies in the same way as they process information in normal melodies. Even when the perceptual groups do not make musical sense, listeners prefer to maintain the information within each group as a complete whole and migrate the click towards the boundary. Again, the results suggest that listeners are anticipating the end of the perceptual group by the time the penultimate note in the group occurs. The second important finding was that listeners respond differently to clicks in a long note the duration of a melody excerpt than they do to clicks in that melody excerpt, even though temporal positions of clicks were identical. Localisation response to the note on A_4 was different to response to melody S2. This suggests that migration effects obtained in the above melodies, especially in melody S2, are due to the musical structure rather than general temporal effects which may have occurred at positions of click placement.

Summary

A set of experiments were carried out to examine localisation response to clicks which were positioned in Gaelic melodies. Given that the task was novel, experiments were performed to check whether listeners were focusing on the musical structure as directed and to examine whether the localisation task was directly affecting responses. Results showed that listeners were in fact attending to the musical structure as directed. A modified localisation task without click detection task yielded similar localisation results as the original task. In line with previous click migration studies, localisation responses to clicks positioned on notes before a main boundary were generally migrated later than their actual occurrence towards the boundary. Similarly, clicks which were positioned on notes after a main boundary were migrated earlier than their actual occurrence towards the boundary. The main experimental results showed consistent migration effects in boundary 1 of three of the melodies, S1, S2 and H1, with a similar trend in responses to melody H2, although migration effects did not reach statistical significance. A detailed examination of migration effects at boundary 2 in melodies S1, S2 and H1 revealed migration in the expected direction only when the boundary did not end the excerpt for two of the melodies. When boundary 2 ended the excerpt, migration occurred in the opposite direction away from boundary. This highlights the effect which experimental stimuli can have on the perceptual responses of the listener. The results suggest that there is a need to control how the musical excerpt is presented to listeners in the experiment. The implications of these studies are that click migration should be studied around boundaries which are not at temporal extremes of the musical excerpt, i.e. not at the beginning or the end of the musical excerpt. Melody S1 did not show the expected migration response around boundary 2, even when the excerpt did not end on boundary 2. This may have been due to the longer overall duration of melody S1 compared with the shorter S2 and H1 excerpts. This emphasises the importance of limiting the excerpts to a memorable length, especially in a task such as the present localisation task which depends upon a memory response from listeners. Present results would suggest that the duration of experimental excerpts does not exceed 10 seconds, although excerpts shorter than this are preferable.

The present study also revealed migration effects in an unstructured melody which were in line with migration effects found in normal melodies. In line with the Segmentation experiment, above, this suggests that the auditory information does not have to make musical sense. Listeners process the information into groups in a similar manner to normal melodies, presumably to reduce the load on memory. Localisation responses to clicks which were positioned within a control stimulus (one note lasting the duration of phrase 1 - 2 of melody S2) demonstrated that the migration effects reported in the present experiments were due to the musical structure rather than general temporal features of the click positions.

The present study shows that this is a useful non-visual technique with which to examine the way in which listeners process information in melodies. Listeners processed the auditory information in the melodies as discrete units. These units corresponded to the perceived phrase structure of the Gaelic melodies which was indicated in the Segmentation Experiment.

Appendix A Mean responses in Condition 1

Table 1 Mean click *localisation* responses to clicks on notes and clicks on middle of notes

	Click positioned *on* note		Click *on middle* of note	
Position of click	mean	s.d.	mean	s.d.
before boundary S2, n = 9	0.06	0.24	-0.07	0.32
after boundary S2, n = 9	-0.27	0.23	-0.22	0.29
before boundary H2, n = 10	0.16	0.25	0.11	0.40
after boundary H2, n = 10	-0.08	0.32	-0.08	0.38

Note " - " response is an early migration

Table 2 Mean click *localisation* responses to clicks positioned before and after three main perceived boundaries

	Melody S2, n = 9				Melody H2, n = 10			
	before boundary		after boundary		before boundary		after boundary	
Boundary number	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.
1	-0.04	0.29	-0.32	0.26	0.00	0.31	-0.28	0.28
2	-0.08	0.39	-0.30	0.31	-0.04	0.38	-0.26	0.32
3	0.13	0.52	-0.12	0.32	0.10	0.51	-0.12	0.30

Note " - " response is an early migration

Table 3 Mean click *detection* responses to isolated clicks, clicks positioned on a note, and clicks positioned on a note in a melody.

	Isolated click		Click on isolated note		Click on note in melody	
	mean	s.d.	mean	s.d.	mean	s.d.
Melody S2, n = 8	0.28	0.09	0.28	0.08	0.34	0.05
Melody H2, n = 8	0.24	0.04	0.31	0.09	0.33	0.08

Table 4 Mean click *detection* responses to clicks on notes and clicks on middle of notes.

	Click on middle of note		Click on note	
	mean	s.d.	mean	s.d.
melody S2, n = 9	0.33	0.12	0.37	0.16
melody H2, n = 10	0.33	0.06	0.33	0.06

Appendix B Mean responses for Condition 2

Table 1a Mean click *localisation* responses to melody H1, *Condition* 2, n = 33

Click position	Mean in seconds (- responses = early migrations)	Standard deviation
1	0.56	0.90
2	0.32	0.34
3	0.25	0.34
4	-0.01	0.26
5	-0.04	0.49
6	-0.18	0.54
7	-0.11	0.53
8	-0.29	0.54
9	-0.35	0.46
10	-0.17	0.49
11	-0.26	0.73
12	-0.47	0.60
13	-0.50	0.47
14	-0.34	0.42
15	-0.41	0.61

Table 1b Mean click *detection* responses to melody H1, *Condition* 2, n = 33

Click position	Mean in seconds	Standard deviation
1	0.50	0.27
2	0.50	0.24
3	0.50	0.25
4	0.37	0.17
5	0.40	0.18
6	0.39	0.19
7	0.33	0.12
8	0.34	0.12
9	0.33	0.11
10	0.32	0.09
11	0.32	0.08
12	0.30	0.10
13	0.32	0.17
14	0.29	0.07
15	0.31	0.10

Table 2a Mean click *localisation* responses to melody S1, *Condition* 2, n = 33

Click position	Mean in seconds (- responses = early migrations)	Standard deviation
1	0.31	0.38
2	0.33	0.43
3	0.36	0.47
4	0.11	0.61
5	0.56	0.87
6	0.00	0.83
7	0.10	0.65
8	-0.71	1.04
9	-0.38	0.85
10	-0.38	0.81
11	-0.62	0.80

12	-1.41	1.04
13	-0.98	1.11
14	-0.57	0.90
15	-0.23	1.13
16	-0.53	0.92
17	-0.30	1.10
18	-0.56	1.14
19	-0.18	1.50
20	-0.84	1.06
21	-0.71	1.10
22	-0.30	0.49
23	0.00	0.66
24	-0.03	0.58
25	-0.23	0.70
26	-0.34	0.66
27	-0.19	0.54
28	-0.55	0.63
29	-0.34	0.70

Table 2b Mean click *detection* responses to melody S1, *Condition 2*, n = 33

Click position	Mean in seconds	Standard deviation
1	0.58	0.24
2	0.52	0.23
3	0.50	0.35
4	0.44	0.20
5	0.41	0.19
6	0.34	0.08
7	0.33	0.09
8	0.36	0.13
9	0.36	0.18
10	0.32	0.09
11	0.31	0.08
12	0.32	0.08
13	0.32	0.06
14	0.34	0.11
15	0.32	0.10
16	0.32	0.07
17	0.31	0.10
18	0.30	0.08
19	0.31	0.08
20	0.32	0.12
21	0.30	0.09
22	0.31	0.13
23	0.35	0.13
24	0.35	0.12
25	0.39	0.14
26	0.34	0.07
27	0.38	0.22
28	0.33	0.12
29	0.36	0.16

Appendix C Mean responses for Condition 3

Table 1 Mean click *detection* responses to melody H1, *Condition 3*, n = 11

Click position	Mean in seconds	Standard deviation
1	0.42	0.10
2	0.42	0.18
3	0.39	0.18
4	0.32	0.10
5	0.28	0.06
6	0.30	0.06
7	0.28	0.09
8	0.33	0.19
9	0.27	0.06
10	0.29	0.06
11	0.31	0.11
12	0.27	0.11
13	0.24	0.10
14	0.27	0.05
15	0.24	0.05

Table 2 Mean click *detection* responses to melody S1, *Condition 3*, n = 11

Click position	Mean in seconds	Standard deviation
1	0.52	0.29
2	0.40	0.15
3	0.47	0.19
4	0.36	0.07
5	0.32	0.08
6	0.27	0.06
7	0.30	0.08
8	0.30	0.09
9	0.30	0.10
10	0.30	0.08
11	0.25	0.05
12	0.28	0.09
13	0.27	0.06
14	0.33	0.11
15	0.29	0.10
16	0.30	0.07
17	0.26	0.06
18	0.24	0.05
19	0.28	0.06
20	0.26	0.12
21	0.26	0.04
22	0.24	0.07
23	0.28	0.15
24	0.37	0.23
25	0.32	0.09
26	0.29	0.07
27	0.27	0.09
28	0.28	0.04
29	0.29	0.10

Appendix D Mean responses for Condition 4

Table 1 Click *localisation* responses to melody H2, *Condition 4*, n = 10

Click position	Mean in seconds (- responses = early migrations)	Standard deviation
1	0.23	0.34
2	0.38	0.42
3	0.43	0.77
4	0.54	0.87
5	0.72	0.87
6	0.46	1.37
7	0.12	0.66
8	0.67	1.12
9	0.23	0.98
10	-0.02	0.35
11	0.08	0.39
12	-0.12	0.43
13	0.10	1.07
14	0.30	0.68
15	-0.20	0.77
16	-0.66	1.32
17	-0.06	0.94
18	-0.24	0.65
19	-0.49	1.06
20	-0.47	0.68
21	-0.54	0.89
22	-0.32	0.74
23	-0.14	0.62
24	-0.49	0.63
25	-0.30	0.58
26	0.01	0.58

Table 2 Click *localisation* responses to melody S2, *Condition 4*, n = 10

Click position	Mean in seconds (- responses = early migrations)	Standard deviation
1	0.37	0.35
2	0.46	0.75
3	0.35	0.75
4	-0.17	0.84
5	0.12	0.58
6	-0.35	0.82
7	-0.69	0.70
8	0.02	0.83
9	0.24	1.08
10	-0.27	0.97
11	-0.57	0.61
12	-0.30	0.88

Appendix E Mean responses for Condition 5

Table 1 Click *localisation* responses to melodies S2 & H2, *Condition 5*, n = 10

Click position	Mean in seconds (- responses = early migrations)	Standard deviation
Melody S2 "ending"	-0.22	0.43
Melody S2 "not ending"	0.07	0.27
Melody H2 "ending"	-0.32	0.31
Melody H2 "not ending"	0.38	0.73

Appendix F Mean responses for Condition 6

Table 1a Click *localisation* responses to melody H1, *Condition 6*, n = 20

Click position	Mean in seconds (- responses = early migrations)	Standard deviation
1	0.13	1.08
2	0.04	0.74
3	-0.25	1.04
4	0.22	1.18
5	0.17	0.48
6	-0.21	0.83
7	-0.44	0.55
8	-0.36	0.64
9	-0.08	0.62

Table 1b Click *detection* responses to melody H1, *Condition 6*, n = 20

Click position	Mean in seconds	Standard deviation
1	0.39	0.14
2	0.36	0.09
3	0.41	0.19
4	0.40	0.16
5	0.36	0.10
6	0.35	0.09
7	0.34	0.08
8	0.38	0.21
9	0.37	0.18

Table 2a Click *localisation* responses to melody S1, *Condition 6*, n = 14

Click position	Mean in seconds (- responses = early migrations)	Standard deviation
1	0.34	0.62
2	0.45	0.49
3	0.69	0.86
4	0.23	0.83
5	0.38	0.68
6	-0.05	1.01
7	0.25	0.82
8	-0.41	1.02
9	-0.68	1.07

10	-0.33	0.94
11	-0.57	0.48
12	-0.59	0.93
13	-0.32	1.31
14	-1.22	1.32
15	-0.25	1.31
16	-0.23	1.69
17	-0.59	1.15
18	-1.21	1.41
19	-0.52	1.40
20	-0.70	1.86
21	-0.66	1.12

Table 2b Click *detection* responses to melody S1, *Condition 6*, n = 14

Click position	Mean in seconds	Standard deviation
1	0.67	0.25
2	0.59	0.22
3	0.56	0.30
4	0.49	0.15
5	0.42	0.07
6	0.43	0.09
7	0.46	0.14
8	0.46	0.17
9	0.39	0.08
10	0.36	0.05
11	0.38	0.09
12	0.37	0.06
13	0.43	0.09
14	0.45	0.16
15	0.34	0.09
16	0.37	0.11
17	0.36	0.07
18	0.34	0.10
19	0.34	0.09
20	0.35	0.08
21	0.37	0.08

Table 3a Click *localisation* responses to melody S2, *Condition 6*, n = 15

Click position	Mean in seconds (- responses = early migrations)	Standard deviation
1	0.48	0.52
2	0.57	0.79
3	0.36	0.69
4	0.30	0.33
5	0.16	0.48
6	-0.47	1.00
7	0.08	0.78
8	0.33	1.35
9	-0.66	1.42
10	0.19	1.25
11	-0.11	0.73
12	-0.33	0.86
13	-0.42	1.23
14	-0.91	1.02

Table 3b Click *detection* responses to melody S2, *Condition 6*, n = 15

Click position	Mean in seconds	Standard deviation
1	.58	.23
2	.56	.19
3	.44	.16
4	.40	.09
5	.39	.12
6	.38	.09
7	.35	.09
8	.36	.09
9	.38	.20
10	.34	.05
11	.36	.11
12	.35	.14
13	.34	.06
14	.42	.36

Appendix G Mean responses for Condition 7

Table 1 Click *localisation* responses to melody Unstructured S2, *Condition 7*, n = 15

Click position	Mean in seconds (- responses = early migrations)	Standard deviation
1	0.12	0.34
2	0.30	0.71
3	0.59	0.69
4	0.43	0.86
5	-0.04	0.53
6	0.32	0.61
7	-0.21	0.55
8	-0.11	1.09
9	-0.34	0.61
10	-0.21	0.69
11	-0.19	0.49
12	-0.03	0.46
13	-0.69	0.64

Table 2 Click *localisation* responses to *One tone* the duration of melody S2 phrase 1 - 2, n = 18

Click position	Mean in seconds (- responses = early migrations)	Standard deviation
1	0.04	0.57
2	-0.10	0.89
3	0.05	0.96
4	0.39	1.09
5	-0.20	1.23
6	0.06	0.70
7	-0.28	1.06
8	-0.43	1.02
9	-1.15	1.21
10	-0.89	0.93
11	-0.85	1.29
12	-1.59	1.76

Chapter 3 Rhythm recall

It has been established that information in Gaelic melodies is consistently chunked into units (Segmentation Experiment, Chapter 2) and demonstrated that listeners process information in these melodies in a way which preserves the integrity of these units (Click localisation Experiment, Chapter 2). The next question of interest is how is information in Gaelic melodies remembered? It has been noted that rhythmic structure plays a major role in determining memory for musical sequences (Chapters 1 & 2). Despite this importance, it has not been given the attention which it deserves in empirical studies of music perception. The present chapter introduces a rhythm recall task which is used to examine how much rhythmic information is remembered from musical sequences and unfamiliar melodies.

Memory for melodies

Deutsch (1980) studied memory for pitches in structurally varying tonal sequences using a melodic dictation task. Musicians heard the experimental sequence and recalled the pitches by writing them on a musical score. Sequences were constructed according to a tight set of musical rules. Temporal segmentation was induced by pauses. The main finding was that tonal structure facilitated recall and where this coincided with compatible temporal segmentation recall performance was optimal. Boltz & Jones (1986) also used a melodic recall task with musicians to examine how rhythmic structures combine with pitch structures to affect memory for musical patterns. Listeners heard a 12 tone melodic sequence which they reproduced on a musical score after one hearing. They found that recall was good for patterns containing a small number of contour changes and for patterns in which accents coincided with points of contour change to make the contour changes more temporally predictable. Jones, Maser & Kidd (1978) also studied pitch recall using a different form of pitch recall whereby listeners either wrote down the pitches they had heard in a row of boxes or identified their occurrence by numbering keys on a keyboard display which corresponded to the perceived pitches. They found that recall was poor for sequences which contained large pitch leaps.

An example of a non-musical task which examined the effects of pitch and rhythm factors on memory is Frankish (1995). He used an immediate recall task to examine memory for 9 digit sequences under a variety of pitch conditions such as familiar melody, unfamiliar melody and monotone. Participants wrote their responses on structured response grids which contained one grid space for each possible serial position. The task was to fill in the gap with the corresponding digit working from left to right. He found that memory for monotone digit sequences was no different for memory for familiar melody digit sequence. Although this is a useful example of a recall task which non-musicians can perform, listeners are not often faced with perceiving digit lists which are paired with musical pitch patterns. In addition to this, the recall task placed additional demands on the listener as they had to remember a series of digits and then write each digit into a corresponding grid space.

Pitch and rhythm features: interrelated

The perceptual effects of pitch and rhythm compatibility in musical sequences have also been reported for different tasks such as phrase completion and pitch discrimination. Boltz (1989) carried out an experiment to examine the effects of the temporal structure of melodies on musicians ratings of melody endings. Listeners rated the completion of a melody on a 10 point scale. Melodies which ended on the leading tone-tonic interval and ended on positions in time which were congruent with the temporal accent structure were judged as being most complete. Schmuckler & Boltz (1994) carried out a similar experiment to examine the effect of temporal and harmonic context on expectancies for melodies containing four chord sequences. Listeners performed a 7 point expectancy or appropriate task to describe how appropriate they perceived the fourth chord to be. Chords were either a traditional cadence, harmonically related or from a key outside the melody. The chords were timed so that they occurred on the first beat of every measure to generate an expectancy of occurrence or were randomly timed. The fourth chord was either early, on time or late in relation to the first beat of the final bar. They found that both temporal and harmonic factors contributed to the perceived appropriateness of the final chord. Harmonically strong cadences which were temporally predictable and occurred on time were judged as the most appropriate. They argue that this demonstrates the need to acknowledge the role of temporal and pitch factors in understanding how listeners mentally represent musical structure. In addition to this, Boltz (1993) used a pitch detection task with folk melodies to study the way in which temporal and pitch accents create expectancies in the listener. She played a melody which contained regular temporal accenting to induce a temporal expectancy in the listener. Listeners were then played two equitemporal tonal sequences and performed a pitch detection task to an altered pitch in the sequence. Results revealed that success on the pitch detection task was related to the preceding rhythmic context. Responses were facilitated in positions where the preceding temporal framework coincided with important pitch events such as phrase endings.

Others have found results which suggest that compatibility of temporal and pitch structures is not the primary factor to influence memory for the pitches of musical sequences. Smith & Cuddy (1989) used a melodic dictation task, similar to that of Boltz & Jones (1986), to study the effects of varying rhythmic structure on listeners memory for tonal sequences. Listeners were required to notate the pitches of the sequence only, all sequences were equitemporal. Sequences were either 'matched' (temporal and pitch accents coincided) or 'mismatched' (temporal and pitch accents did not coincide). They report that pitch recall accuracy was affected primarily by characteristics of the metric rhythm rather than the compatibility of temporal and pitch patterns.

Pitch and rhythm: independent

This lack of effect of compatibility of temporal and pitch structures on perceptual responses has been pushed further by other researchers who report that melodic and rhythmic patterns have independent effects on listeners' performance in tasks such as similarity ratings for pairs of musical segments and phrase completion. Monahan & Carterette (1985) studied musicians similarity ratings of melody pairs which varied along rhythm or pitch or both rhythm and pitch dimensions. A same-different task was utilised. Multidimensional analysis showed that at least five different dimensions were needed to describe the perceptual responses to the melodies. The first three dimensions were interpreted as rhythmic which highlights the importance of rhythmic structure on perceptual responses to musical stimuli. The last two dimensions were related to pitch features in the melodies. They found that listeners who rated the melodies along rhythm dimensions did not rate along pitch dimensions and vice versa. This has been taken as evidence for the independence of pitch and rhythm features in melody perception. Palmer & Krumhansl (1987a; 1987b) also report an independent relationship between pitch and rhythm features in phrase perception of a Mozart piano sonata and a melody by Bach. Listeners heard the original music, the pitches alone and the rhythm alone and performed a goodness/completion of phrase rating task. They found that listeners responses to the complete phrase were highly correlated with both the rhythm and pitch alone conditions. The pitch and rhythm alone conditions were not correlated with each other which led researchers to conclude that phrase perception is the result of separate pitch and temporal patterns.

Pitch and rhythm: a compromise

Neuropsychological evidence offers some form of compromise between the two opposed view of pitch and rhythm as integrated or independent structures. Peretz & Kolinsky (1993) found evidence to support the view that melody and rhythm are treated as independent dimensions which are integrated into a unified percept at a later stage of processing. Their evidence came from a patient with bilateral temporal lobe damage whose performance on a same-different task revealed that they processed melodic and temporal information independently of each other. Although their patient showed that melody and rhythm can be processed separately, Peretz & Kolinsky (1993) examined normals and found evidence to suggest that rhythm and melody are integrated. Performance on a same-different rhythmic dimension was affected by the pitch dimension. At what level in the processing stage rhythm and pitch factors are integrated remains unclear. It is important to note that when listeners were required to concentrate solely upon the rhythm, the effect of pitch on task performance was reduced. This suggests that the effect of pitch on rhythm discrimination may be affected by the listeners attentional focus at a given point in time. This is an important issue for all tasks which examine pitch and rhythm compatibility on pitch and rhythm perception.

Theoretical approaches

Thompson (1994) also offers a way of resolving the issue of separate pitch and rhythm structures using a feature-integration theory originally used in visual processing research. The basic idea is that an initial stage of feature detection occurs whereby musical features such as pitch, rhythm and timbre are analysed separately. This is followed by an integration phase in which, for example, pitch and rhythmic features are integrated most successfully under optimal processing conditions such as compatible pitch and temporal patterns. Thompson (1994) examined listeners' ability to detect changes in the way pitch and duration are combined in individual note-note pairs and short melodies. Stimuli were altered in pitch, in duration or in both pitch and rhythm. A test stimuli was played, then repeated and listeners indicated whether it had changed or not. This was performed under an attentive condition or a non-attentive condition in which the listener performed a word description task throughout the stimulus presentation. Thompson (1994) found that attentive listeners detected changes in both parameters for note pairs and short melodies. Non-attentive listeners could not detect changes in both parameters successfully. It was concluded that pitch and duration parameters are processed independently at a pre-attentive stage. Thompson suggests that the integration of these musical parameters is influenced by focused attention at the level of perception and short term memory.

A recent study also highlights the importance of type of musical stimuli in addition to focus of attention on the results obtained in the only rhythm recall experiment to date. Rhythm recall of unfamiliar folk tunes was examined using a letter reproduction task under a number of attending conditions. Listeners were asked to attend to the pitch only, the rhythm only or both the pitch and rhythm of a melody (Boltz, 1997). Melodies consisted of three different events durations which were recalled by writing down L (long), M (medium) and S (short). Melodies were also classed as coherent: pitch and temporal accents coincided with one another and incoherent: pitch relations and temporal accents conflicted with melodic phrase boundaries. Boltz (1997) found that rhythm recall of coherent melodies was highly successful for all attending conditions. Recall of incoherent melodies was generally poorer, with the worst recall for both attending then for pitch attending. This suggests that pitch and rhythmic information is jointly encoded in coherent melodies and encoded separately for incoherent melodies.

Despite the contrasting findings of independency of pitch and temporal features in music perception, there is a large body of evidence which has shown that pitch and rhythm factors interact in the perception and memory of musical sequences. It has been noted that an examination of whether there are independent or integrated effects of pitch and rhythm may not be informative because opposing results may due to different task demands and the focus of attention during responding (Peretz & Kolinsky, 1993). An example is the recall task on which most researchers found an interdependent effect of pitch and rhythm features (Deutsch, 1980; Boltz & Jones, 1986) and a same-different task which found independent effects of pitch and rhythmic features (Monahan & Carterette, 1985). It is possible that pitch recall uses both pitch and rhythm structures as the musician sings the sequence back in their head whilst writing down the pitches. The same-different task may encourage the listener to focus upon a salient feature of the melody and examine whether this is repeated in the comparison melody. Listeners may develop a strategy whereby they notice, for example, a particularly jumpy rhythm with which to compare and contrast the rhythm of the second fragment they hear. It may be easier to focus on one structure or the other in this task and results may be due to which structure is being consciously focused upon (Peretz & Kolinsky, 1993).

It has also been pointed out that most research in this area has examined memory for *pitch* features whilst varying the rhythmic structure of the sequence (discussion above). What about memory for *rhythm*? In addition to this there are many criticisms directed towards the stimuli used in this type of research. There are problems with studying pitch effects (e.g. pitch contour effects) on memory in equitemporal rhythmic sequences as music does not typically rely on such equitemporal rhythms. It is often the case that a small number of experimental sequences are used to examine perceptual responses. There is a strong and interrelated patterning of temporal and melodic events in real music which is often ignored in the simplistic experimental stimuli predominantly used in the majority of memory experiments which examine pitch recall (Jones, 1987). As a result, experimental stimuli are often too impoverished to lead to any firm conclusions as to the possible effects of compatible pitch and rhythmic structures on memory for real music. These two issues can be directly addressed. Memory for the *rhythmic* structure of *real* melodies which contain repetitive and meaningful pitch and rhythm structures should be examined. The questions of interest are:

1. Will memory for rhythmic structure be affected by pitch structure in the same was as memory for pitch structure is affected by temporal structure (Boltz, 1986; Dowling, 1973)?

2. What kind of musical units will listeners remember from the Gaelic melodies? Do these units correspond to the information units identified in the Segmentation experiment of chapter 2?

The next section introduces the task used to examine non-musicians memory for rhythm.

Rhythm Reproduction

A number of studies have examined the ability of listeners to reproduce musical rhythms. Although this type of task clearly involves some form of memory component, the primary focus is on basic rhythm performance abilities and how performance relates to the music. Sequences are typically repeated so that listeners can remember the structure. Researchers examine which type of rhythmic structure is easier to perform. Drake (1993) studied the reproduction of musical rhythms by children, novices and musicians. A motor response task was used in which participants tapped back the perceived rhythm on a drum. Drake found that rhythms containing durations in the ratio of 1:2 were easier to reproduce than rhythm in a ratio of 1:3. Sequences which contained a fewer number of different durations were more successfully reproduced, for example, a crotchet and quaver sequence was easier to beat back than a crotchet, quaver and semiquaver sequence. Gerard & Auxiette (1988) examined the effects of verbal and melodic cues on the reproduction of musical rhythms by children. They used a clapping task to measure rhythm reproduction. They found that rhythm reproduction was facilitated by verbal and melodic cues in children who had previous experience with such cues. Other experiments examining reproduction have been interested in the accuracy with which musicians reproduce rhythmic durations of complex rhythmic patterns (Gabrielsson, 1974).

It has been established that rhythmic structure, albeit basic rhythmic structure, has an effect on pitch memory (Dowling, 1973; Deutsch, 1980; Boltz & Jones, 1986). There is a need to examine whether memory for the rhythms of real melodies is affected by pitch features in those melodies. As discussed above, Boltz (1997) used a letter reproduction task to examine rhythm recall. This task can be used successfully with non-musicians. However the task demands are unnatural. A lot of effort is required to process musical information and perform a recall task which requires the transformation of perceived rhythmic structures into verbal symbols which are then written down. A more suitable recall task is the simple tapping back of the perceived rhythmic structure. Rhythm reproduction tasks have been used successfully from a performance analysis point of view. Rhythm reproduction is an ideal task with which to examine memory for rhythmic structure. Non-musicians can easily perform this task immediately after hearing the musical excerpt.

The present research

This research was conducted to investigate short-term memory (STM) for a variety of musical note sequences and Gaelic melodies. The aim was to enhance understanding of the factors which facilitate or disrupt STM for musical sequences whilst addressing the imbalance of the large amount of pitch recall research in comparison to the almost non-existent rhythm recall work. The interest lay in examining how much rhythmic information listeners can extract after one hearing of an unfamiliar melody. The effects of altering the pitch pattern of a musical sequence on STM for the rhythm of that sequence was also examined. The task used to assess recall performance was a rhythm reproduction task which required listeners to tap back as accurately as possible the rhythm they perceive. Listeners were informed that they were participating in a rhythm perception experiment with no mention of the memory component of the task. This was to avoid listeners developing unusual strategies for remembering the sequences which may disrupt the normal listening experience. No specific instructions were given about ignoring or attending the pitch pattern which accompanied the melodies. Listeners were informed that their task was to tap back the rhythm as accurately as possible; the corresponding pitch patterns were not mentioned. The following two experiments were carried out:

Experiment I which examined rhythm recall of computer generated melodies ranging in length from 3 - 7 beats. The melodies contained pitches which were either *structured* (in the key of C major with small pitch intervals of 1 - 4 semitones) or *unstructured* (contained any notes from the octaves immediately above and below C_4 with large pitch intervals of 4 semitones up to 2 octaves). The rules governing the construction of these melodies are detailed below in the methods section, pp 105 - 106.

Experiment II which examined rhythm recall of Gaelic melodies under a variety of structural alterations. These alterations are described in more detail in the methods section, page 117.

Experimental hypotheses

The hypothesis for Experiment I is that the atonal sequences with large pitch intervals (unstructured) will disrupt rhythm recall if pitch and rhythmic factors are perceived jointly. The hypotheses for Experiment II are: rhythm recall performance will be optimal for sequences with repetitive and meaningful rhythmic and pitch structure (original melody). Secondly, sequences with the greatest degree of structural alteration will be the hardest to recall (both rhythm and pitch reversed).

Experiment I

Rhythm recall of tonal sequences containing small intervals vs atonal sequences containing large intervals

This experiment was carried out to examine the effects of pitch-interval structure on rhythm recall by comparing performance of tonal with small interval and atonal with large interval sequences. An objective measure of the amount of sequence which is retained in memory was obtained from a pilot study which examined recall of structured sequences ranging from 2 - 13 beats in length. This established that recall performance decreases markedly after 7 beat length sequences and that listeners are consistently remembering an average of around 3 beats worth of the sequence. Another pilot study revealed that presentation rate did not effect recall performance to any significant extent. The rates ranged from *slow*: one crotchet beat per 2 seconds, *slow/medium*: one crotchet beat per second, *medium*: one crotchet beat per 0.60 second and *fast*: one crotchet beat per 0.30 second. Increasing the presentation rate did not enable the listener to retain more of the sequence in memory. The present experiment used the medium tempo of one crotchet beat per 0.60 seconds as the rate of presentation.

Method

Subjects The subjects were 31 undergraduates, 18 males, 13 females, aged between 18 and 29 years. Participation in the experiment was in part fulfilment of a psychology course requirement. All subjects were naive as to the purposes of the experiment.

Stimuli Test sequences were produced on a Macintosh Centris 650 using a pipe organ timbre and were delivered binaurally through headphones. The rhythms were constructed with the following constraints:

1. All first and last notes were crotchets (crotchets, quavers and semi-quavers were the note values which could have occurred within each given sequence).
2. The beats were always regular: there were no rests, dotted or tied notes and consequently no syncopations.
3. The pattern of note durations within each sequence was created according to a set of probabilities which determined whether note durations increased, decreased or remained the same. The probabilities were adjusted through a sequence to create a musically acceptable rhythm. The algorithm used to generate the rhythm sequences is displayed in Appendix 1, page 144.

The pitch patterns of the tonal sequences with small pitch intervals (*structured*) were constructed according to the following constraints:

1. All notes were chosen from the scale of C Major.
2. All first notes were C_4.
3. All subsequent pitches were chosen at random to be 1 or 2 scale steps away from the present one (with equal probability). The probability of a pitch interval being 1, 2, 3, or 4 semitones in width was 0.15, 0.39, 0.23 and 0.23, respectively. No pitch intervals greater than 4 semitones occurred in the tonal sequences. The direction, up or down, was determined separately. The first interval always moved upwards after which the direction has a random probability of 0.25 of changing direction and 0.75 of maintaining the same direction. When these rules took the pitch beyond the octave above or below C_4, the direction was altered to bring the next note back into the octave above or below C_4. The tempo was one crotchet beat per 0.60 seconds and all sequences were unaccented. Listeners heard the sequences at comfortable listening levels. These tonal sequences were characterised by tonal runs of small pitch intervals (1 - 4 semitones) in the same direction. The label *structured* will refer specifically to these sequences in the remainder of this thesis.

The pitch pattern probability set was altered for the sequences with large pitch intervals (*unstructured*) to create a random sequence of pitches from the octaves immediately above and below C_4. The sequences were created with the following constraints:

1. All first notes were C_4

The succeeding notes in the sequence were selected from the twelve notes of the chromatic scale in the octave above C_4 and the twelve notes of the chromatic scale in the octave below C_4. The probability of a pitch interval 1, 2, 3 or 4 semitones apart was 0.08, 0.077, 0.073 and 0.07, respectively. The probability of occurrence of pitch intervals greater than 4 semitones apart was 0.70. These atonal sequences were characterised by erratic pitch interval leaps which ranged from 1 - 24 semitones and favoured pitch intervals greater than 4 semitones apart. The label *unstructured* will refer specifically to these sequences in the remainder of thesis.

Procedure Listeners were run individually in the presence of the experimenter in a quiet room. They were informed that they were would hear a series of tone sequences and were asked to tap back as accurately as possible the rhythm which had been presented after one hearing. The experimenter demonstrated the task by running as participant for three trials. Any key could be used to tap the rhythm and listeners were requested to use the key with which they felt most comfortable. After each sequence listeners were given feedback on the computer screen indicating percent correct for their performance. Listeners were instructed to proceed at their own pace. The visual prompt "Ready?" appeared on screen at the start of each sequence, followed by the instructions "Please listen to the Rhythm" as the sequence ran and a final "Please tap the rhythm" when the sequence ended. Five practise trials were run before the experimental trials began. There were two blocks, one structured the other unstructured pitches, of 45 trials. Listeners performed on both blocks and presentation order was counterbalanced across subjects. Performance on practise trials was discarded. The duration of the experiment was an average of 30 minutes including pauses.

Analysis

As discussed above the main interest lay in examining how much of the rhythm of the sequence listeners are remembering after one hearing. Subjective performance of the rhythm of each sequence must be compared to the objective rhythm of the experimental sequence to examine at which point errors occurred and memory is inaccurate. There are important factors which must be taken into account when choosing the level of accuracy for the purposes of performance analysis. The basic problem is to avoid using an accuracy level which is too lenient and thus results in high scores on every rhythm performance. It should also be noted, however, that the primary interest lay in examining how much of the sequence is remembered, rather than how well listeners could perform the musical rhythms of each melody. Accuracy levels which are too strict could result in low recall scores for every performance. This problem was dealt with by using three different accuracy levels with which to mark each rhythm performance: 80, 70 and 60 %. These are discussed in more detail below.

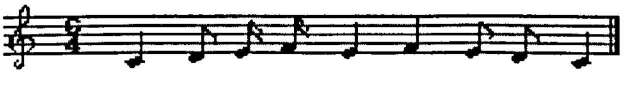
The proportion of the sequence recalled correctly was measured in notes and beats. The raw inter-onset-intervals of the performance were compared with the inter-onset intervals of the corresponding stimuli to give a measure of number of events correct. A ratio-based analysis was utilised. Each event duration in the stimuli sequence was divided by the event duration directly preceding it to give a series of ratios corresponding to the actual durations of the sequence. For response patterns each inter-onset-interval was divided by the inter-onset-interval directly preceding it, thus a corresponding set of two ratio sequences could be compared and scored (see Figure 1, below, for a detailed example of this). The following two methods of scoring were used:

Score I Marked the number of events correctly recalled up to the first single error only

Score II Marked the number of events correctly recalled up to the occurrence of two errors.

The example of Score 1 and Score 2 shown below in Figure 1, below, give the same recall score but this is not necessarily the case. If any differences arise, Score II, being the less stringent scoring method, yields a higher number of events recalled compared with Score I.

Figure I Sample 6 beat structured sequence with 70 % accurate scored response



Stimuli durations	8		4		2		2		8		8		4		4		8
Response durations	32		16		9		9		40		29						
Stimuli ratios		0.5		0.5		1		4		1		0.5		1		2	
Response ratios		0.5		0.56		1		4.44		0.72							
Score I		1		1		1		1		1		0		0		0	
Score II		1		1		1		1		1		0		0		0	
Score notes	6 (correct ratios + 1)																
Score crotchet beats	4 (total correct stimuli durations (8 + 4 + 2 + 2 + 8 + 8 = 32) / crotchet beat duration (8))																

The analysis system also marks responses to a stipulated degree of accuracy. The data were analysed separately to accuracies of 80 %, 70 % and 60 %. This meant that any inter-onset-interval response ratio which fell outside e.g. +/- 20 % of the stimuli inter-onset-interval ratio set would be marked as incorrect at the 80 % accuracy level. Exactly the same process was used for 70 % and 60 % analyses with responses outside +/- 30 % and 40 % of the stimuli inter-onset interval ratio set marked as inaccurate at the 70 % and 60 % accuracy levels. The lower the accuracy level, the higher the recall scores. An example will clarify the accuracy level process of analyses. Take the example in Figure 1, above. The first ratio in the stimulus is 0.5. The first response ratio is also 0.5 which would be marked as accurate on all response levels. The second stimuli ratio is 0.5 and the second response ratio is 0.56. 20 % of 0.5 is 0.1, so any response out with +/- 20 % (i.e. less than 0.4 or greater than 0.6) is marked as an incorrect response at the 80 % accuracy level. The response ratio above is 0.56 so it is considered correct within an 80 %, 70 % and 60% accuracy level. The same process is used for 70 % and 60 % accuracy levels except that the percentage used to determine whether the responses are inaccurate or not is 30 % and 40 %, respectively. The mean results of each accuracy level analysis for the present data are displayed in Figures 1a - 3b, page 112 - 113, below.

Results

One of the main considerations of the present data analysis lies in finding the unit which best represents the amount of rhythmical information remembered by the listener. Information remembered can be considered as units of time in seconds, or number of notes in a melody or beats e.g. crotchet beats recalled. The memory element which underlies the present recall paradigm necessitates a unit which measures the amount of information recalled in time. This unit must be consistent and meaningful for comparative purposes across different stimuli. As has been described above, the present range of stimuli were played at a rate of one crotchet beat per 0.60 seconds. The beat is a consistent unit of time and a natural unit of musical rhythm (Dowling and Harwood, 1986). Analysing the amount of information recalled in terms of how many crotchet beats are recalled also allows us to consider how much information in seconds listeners are capable of recalling. It is also important to know the number of events listeners are capable of handling in memory. Considering recall in terms of time or beats alone does not account for the specific patterns in the number of events recalled in a given time period. An analysis purely by notes, however, would seriously undermine the temporal element involved in recalling rhythmical events e.g. recalling four quavers, two semiquavers would be scored as the same as recalling four crochets and two quavers even though the later is double the temporal load than the former. It is clearly important to take account of the number of notes *and* beats which are recalled by listeners.

An analysis using notes and beats was performed with the amount of the sequence recalled given in notes and beats. A series of Spearman's correlations were performed on the data to examine the relationship between the two different scoring methods. Responses were grouped across subjects and took beat versus note analysis, method of scoring data (either I: marked up to the first error only or II: allowing single isolated errors) and accuracy level (80 %, 70 % or 60 %) into account.

Table 1 Beat versus note correlations

Score I

variable	notes 80 %	notes 70 %	notes 60 %
beats 80 %	0.758		
beats 70%	-------	0.888	
beats 60 %	-------	-------	0.890
Score II			
beats 80 %	0.862		
beats 70 %	-------	0.844	
beats 60 %	-------	-------	0.776

All $p < 0.001$

Table 2 Scoring method: I stringent vs II non-stringent

Analysis by Beats

	Score I		
Score II	80 %	70 %	60%
beat 80 %	0.941		
beat 70 %	-------	0.956	
beat 60 %	-------	-------	0.927

All $p < 0.001$

Analysis by notes

	Score I		
Score II	80 %	70 %	60 %
note 80 %	0.815		
note 70 %	-------	0.906	
note 60 %	-------	-------	0.783

The pattern of results are similar for analysis by notes and by beats (Table 1, above). The pattern of results for scoring method (either stringent I or non-stringent II) are also very similar (Table 2, above). The more stringent Score I was used for further statistical investigation of the data as it was a simpler measure of the amount of sequence recalled by the listener. A further series of correlations were carried out to examine the relationship between each of the accuracy levels on which the data were graded (either 80 %, 70 % or 60 %). See Table 3, below.

Table 3 Accuracy level correlations

Analysis by Beats

Accuracy	80 %	70 %	60 %
80 %	---		
70 %	0.961	---	
60 %	0.921	0.980	---

All $p < 0.0001$

Analysis by notes

Accuracy	80 %	70 %	60 %
80 %	---		
70 %	0.909	---	
60 %	0.840	0.962	---

The pattern of responses scored at the three separate accuracy levels are similar with strong positive correlations. Accuracy level is included as a factor in the detailed statistical analyses below.

Detailed Statistical Analysis on Recall Performance

A series of repeated-measures ANOVA's were performed to establish the effect of pitch interval structure on rhythm recall performance. The within-subjects factor was pitch interval structure (either *structured*: small tonal runs or *unstructured*: random chromatic pitch patterns). Accuracy level (80 %, 70 % & 60 %) was the between-subjects factor. Recall performance means, in notes and beats, are displayed in Figures 1a, 1b, 2a, 2b, 3a & 3b, below.

Number of notes recalled

3 beat sequences

There was a significant effect of pitch pattern on rhythm recall $F(1, 90) = 53.72$, $p < 0.001$, a significant effect of accuracy level $F(2, 90) = 12.14$, $p < 0.001$ and no interaction, $p > 0.90$. Rhythm recall of structured sequences is more successful than unstructured sequences.

4 beat sequences

These results were repeated with a significant effect of pitch pattern on rhythm recall, $F(1, 90) = 18.64$, $p < 0.001$, an effect of accuracy level $F(2, 90) = 12.30$, $p < 0.001$ and no significant pitch pattern - accuracy level interaction, $p > 0.80$.

5 beat sequences

As above, significantly more was recalled from structured sequences than unstructured sequences, $F(1, 90) = 13.11$, $p < 0.001$. There was a significant accuracy level effect $F(2, 90) = 10.63$, $p < 0.001$ and no interaction, $p > 0.30$.

6 beat sequences

Again, recall of structured sequences was better, $F(1, 90) = 13.82$, $p < 0.001$, there was a significant effect of accuracy level $F(2, 90) = 6.96$, $p < 0.005$ and no interaction, $p > 0.90$.

7 beat sequences

Recall was significantly higher for structured compared with unstructured sequences, $F(1, 90) = 10.11$, $p < 0.005$, with a significant accuracy level effect, $F(2, 90) = 12.89$, $p < 0.001$ and no interaction, $p > 0.50$.

In summary, the number of notes recalled was significantly greater for the tonal sequences which contained mainly small pitch intervals than for the atonal sequences which contained mainly large pitch intervals. The significant difference between accuracy levels was not surprising; the more lenient the scoring level, the more recalled. The main point to note is that there were no significant interactions between pitch-interval pattern and accuracy level. Thus the significant effect of pitch structure was consistent for all three accuracy levels.

Number of beats correctly recalled

3 beat sequences

There was a significant effect of pitch pattern on rhythm recall performance, $F(1, 90) = 38.26$, $p < 0.001$, an effect of accuracy level, $F(2, 90) = 11.12$, $p < 0.001$ and no pitch pattern - accuracy level interaction, $p > 0.90$. A greater number of beats were recalled from the structured sequences compared with unstructured sequences.

4 beat sequences

The significant effect of pitch pattern was replicated in the longer 4 beat sequences, $F(1, 90) = 9.22$, $p < 0.005$, with an effect of accuracy level, $F(2, 90) = 7.03$, $p < 0.001$ and no interaction, $p > 0.90$.

5 beat sequences

Again, significantly more was recalled from the structured sequences compared with unstructured, $F(1, 90) = 6.16$, $p < 0.05$, with an effect of accuracy level, $F(2, 90) = 6.09$, $p < 0.005$ and no interaction, $p > 0.40$.

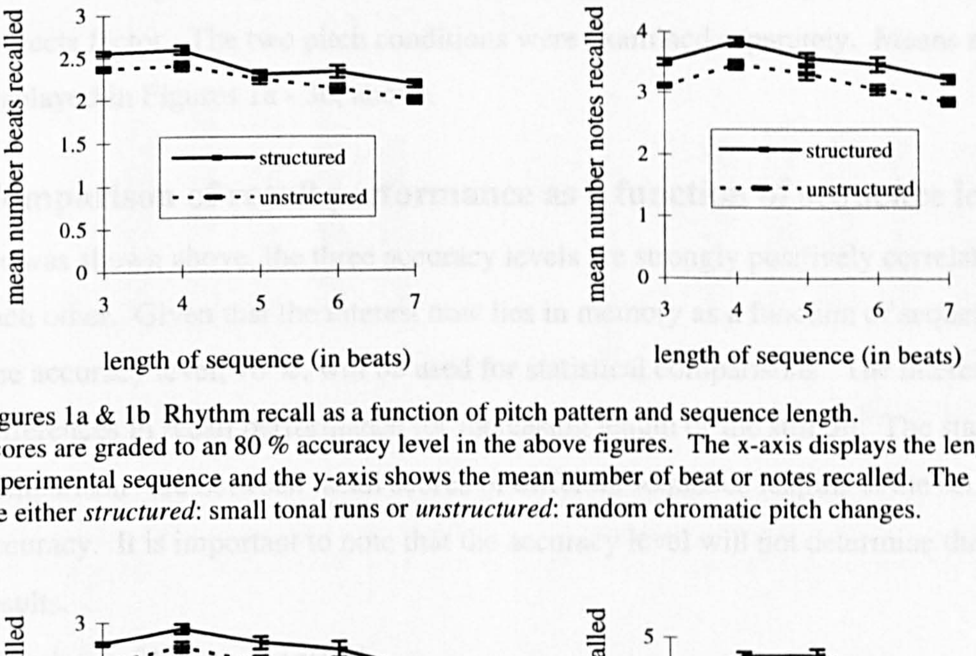
6 beat sequences

Rhythm recall was facilitated for structured sequences in comparison to unstructured sequences, F(1, 90) = 10.88, $p < 0.001$, with an effect of accuracy level, F(2, 90) = 4.47, $p < 0.05$ and no interaction, $p > 0.80$.

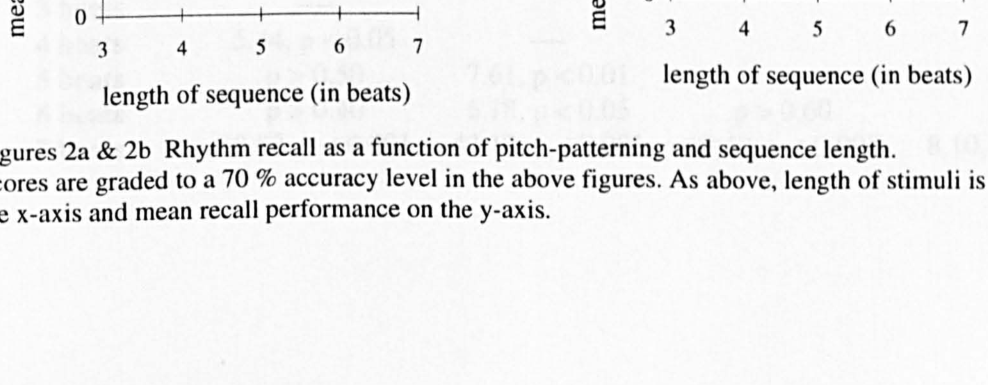
7 beat sequences

Again, there was a significant effect of pitch pattern, F(1, 90) = 5.71, $p < 0.05$, an effect of accuracy level, F(2, 90) = 8.10, $p < 0.001$ and no interaction, $p > 0.70$.

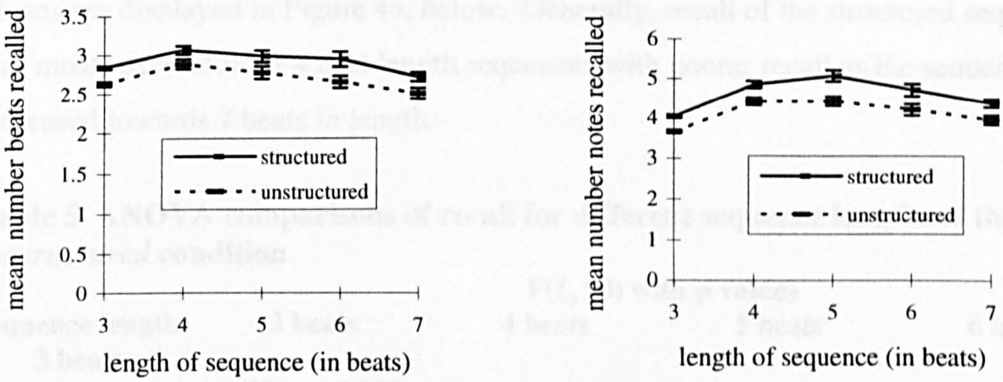
In summary, the beat analysis also revealed that significantly more beats were recalled from tonal with small interval sequences than from atonal with large interval sequences. This effect was consistent across all sequence lengths. Again, there were no significant pitch pattern - accuracy level interactions.



Figures 1a & 1b Rhythm recall as a function of pitch pattern and sequence length. Scores are graded to an 80 % accuracy level in the above figures. The x-axis displays the length of the experimental sequence and the y-axis shows the mean number of beat or notes recalled. The sequences are either *structured*: small tonal runs or *unstructured*: random chromatic pitch changes.



Figures 2a & 2b Rhythm recall as a function of pitch-patterning and sequence length. Scores are graded to a 70 % accuracy level in the above figures. As above, length of stimuli is denoted by the x-axis and mean recall performance on the y-axis.



Figures 3a & 3b Rhythm recall as a function of pitch-patterning and sequence length. Scores are graded to a 60 % accuracy level in the above figures. As above, length of stimuli is denoted by the x-axis and mean recall performance on the y-axis.

A further series of repeated measures ANOVA's examined the differences between rhythm recall performance as a function of the overall length of the sequence (3, 4, 5, 6, or 7 beats long). Length was the within-subjects factor with subjects as the between-subjects factor. The two pitch conditions were examined separately. Means are displayed in Figures 1a - 3b, above.

Comparison of recall performance as a function of sequence length

As was shown above, the three accuracy levels are strongly positively correlated with each other. Given that the interest now lies in memory as a function of sequence length, one accuracy level, 70 %, will be used for statistical comparisons. The interest lies in differences in recall performance for increasing length of the stimuli. The statistical comparisons are between recall scores of different sequence lengths at the same level of accuracy. It is important to note that the accuracy level will not determine the statistical results.

Number of beats recalled

Table 4 ANOVA comparisons of recall for different sequence lengths of the *structured* sequence

	F(1, 90) with p values			
sequence length	**3 beats**	**4 beats**	**5 beats**	**6 beats**
3 beats	----			
4 beats	5.44, $p < 0.05$	----		
5 beats	$p > 0.50$	7.61, $p < 0.01$	----	
6 beats	$p > 0.40$	6.18, $p < 0.05$	$p > 0.60$	----
7 beats	18.82, $p < 0.001$	42.10, $p < 0.001$	10.40, $p < 0.005$	8.10, $p < 0.005$

Means are displayed in Figure 4a, below. Generally, recall of the structured sequences was most successful for 4 beat length sequences with poorer recall as the sequences increased towards 7 beats in length.

Table 5 ANOVA comparisons of recall for different sequence lengths of the *unstructured* condition

	F(1, 90) with p values			
sequence length	**3 beats**	**4 beats**	**5 beats**	**6 beats**
3 beats	----			
4 beats	8.53, $p < 0.005$	----		
5 beats	$p > 0.90$	4.96, $p < 0.05$	----	
6 beats	$p > 0.10$	18.96, $p < 0.001$	$p > 0.09$	----
7 beats	11.40, $p < 0.001$	29.97, $p < 0.005$	16.11, $p < 0.001$	$p > 0.08$

Means are displayed in Figure 5a, below. To summarise, recall of unstructured sequences was also optimal for 4 beat sequences. Performance decreased steadily as sequence length increased towards 7 beats.

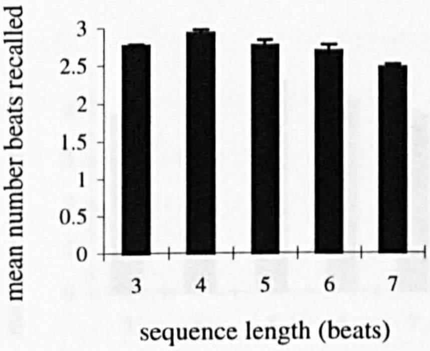


Figure 4a Effect of sequence length on recall of *structured* sequences at 70 % accuracy level.

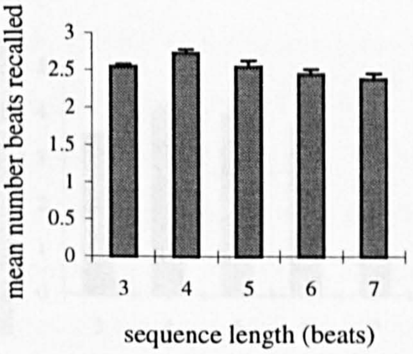


Figure 5a Effect of sequence length on recall of *unstructured* sequences at 70 % accuracy level.

The mean number of beats recalled are displayed on the y-axis, while the sequence length is shown on the x-axis.

Number of Notes recalled

Table 6 ANOVA comparisons of recall for different sequence lengths of the *structured* sequence

	F(1, 90) with p values			
sequence length	**3 beats**	**4 beats**	**5 beats**	**6 beats**
3 beats	----			
4 beats	42.64, $p < 0.001$	----		
5 beats	18.61, $p < 0.001$	$p > 0.90$	----	
6 beats	$p > 0.05$	4.50, $p < 0.05$	5.46, $p < 0.05$	----
7 beats	$p > 0.70$	56.93, $p < 0.001$	26.66, $p < 0.001$	6.69, $p < 0.05$

Means are shown below in Figure 4b. Recall of the structured sequence was most successful for the 4 and 5 beat length sequences. Performance tended to decrease as sequence length increased towards 6 and 7 beats.

Table 7 ANOVA comparisons of recall for different sequence lengths of the *unstructured* sequence

	F(1, 90) with p values			
sequence length	**3 beats**	**4 beats**	**5 beats**	**6 beats**
3 beats	----			
4 beats	44.48, $p < 0.001$	----		
5 beats	15.69, $p < 0.001$	$p > 0.40$	----	
6 beats	4.41, $p < 0.05$	6.35, $p < 0.05$	$p > 0.10$	----
7 beats	$p > 0.40$	15.77, $p < 0.001$	26.66, $p < 0.001$	$p > 0.10$

Means are displayed in Figure 5b, below. A similar pattern of responses emerged for the unstructured sequences with the most successful recall of 4 and 5 beat length sequences and poorest recall of 7 beat sequences.

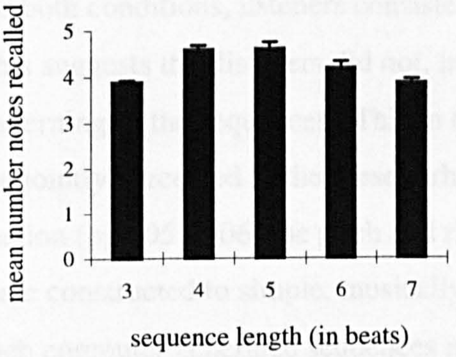


Figure 4b Effect of sequence length on recall of *structured* sequences at 70 % accuracy level.

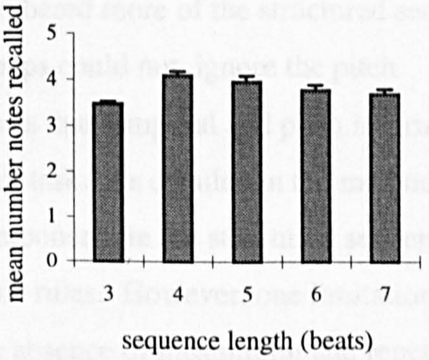


Figure 5b Effect of sequence length on recall of *unstructured* sequences at 70 % accuracy level.

The mean number of notes recalled are displayed on the y-axis, while the sequence length is shown on the x-axis

Discussion

The results of the present recall experiment show that listeners are consistently remembering around 3 beats or 4 - 5 notes worth of rhythmic information from the tonal sequences containing small pitch intervals. Given the tempo of the sequences (one crotchet beat per 0.60 seconds), this is equivalent to retaining 2 seconds worth of information in memory. Thus it could be argued that the present findings are evidence for a constant STM capacity when considered as a unit of time. Disregarding three beat sequences (as the maximum number of beats recalled is three), the mean number of beats recalled decreased as a function of increasing sequence length for both structured and unstructured conditions. This decrease, although relatively small, was more pronounced for structured than for unstructured sequences. A significantly higher number of beats were recalled from the 4 beat sequences for structured and unstructured sequences compared to all other sequence lengths. Recall of 6 beat compared with 7 beat *unstructured* sequences was similar but recall of 6 beat and 7 beat *structured* sequences was significantly different. The present results suggest that there is an overall decrement in performance with increasing sequence length which is more notable for structured sequences than unstructured sequences.

The most important finding of this experiment is the significant difference between recall on the two different pitch conditions. Tonal sequences with small pitch intervals were recalled more successfully than atonal sequences with large pitch intervals for all sequence lengths. Despite the fact that the rhythm patterns to be recalled were the same for both conditions, listeners consistently remembered more of the structured sequences. This suggests that listeners did not, indeed perhaps could not, ignore the pitch patterning of the sequences. This in turn suggests that temporal and pitch information are jointly perceived in the present rhythm recall task. As detailed in the methods section (pp 105 - 106) the pitch and rhythm components in the structured sequences were constructed to simple, musically acceptable rules. However, one limitation of such computer generated sequences may be the absence of meaningful and repetitive pitch and rhythm structures which are prevalent in music. Experiment II examined the same rhythm recall response to Gaelic melodies in a variety of structurally altered conditions.

Experiment II Rhythm recall of Gaelic melodies

Condition 1 Effect of pitch and rhythm alteration on rhythm recall

This experimental condition was carried out to examine the effects of alterations to the musical structure of the melodies on rhythm recall. It was hypothesised that recall of the original melodies, which contain repetitive rhythm and pitch structures, would be most successful. Furthermore, that recall of sequences with the greatest degree of structural alteration (reversal of both rhythm and pitch structures) would be poorest. These are discussed in more detail, below.

Method

Subjects The subjects were 62 undergraduates, 23 males and 39 females, aged between 18 and 38 years. Participation in the experiment was in part fulfilment of a psychology course requirement. All subjects were naive as to the purposes of the experiment.

Stimuli A variety of excerpts from Gaelic melodies were used as experimental stimuli. These were generated using a program on a Macintosh Centris 650 and delivered binaurally through headphones. A piano timbre was used to generate the notes. There were four main conditions for this experiment:

1. Original melody: an excerpt from a Gaelic melody was used.
2. Rhythm reversed: the rhythm of the original melody excerpt was reversed, pitch patterns of the original are preserved.
3. Pitch reversed: the pitch of the original melody excerpt is reversed, the rhythmic structure of the original melody excerpt is preserved.
4. Both reversed: the rhythmic structure and pitch structure of the original melody excerpt is reversed.

The stimuli used in the present experiment are displayed in Figure IIa, Appendix 1, pp 137 - 138. All conditions contained the same number of pitches and event durations, so were temporally comparable and contained equal numbers of elements to be remembered. The differences arose in the degree of structural alteration which was applied to the melody. The two main hypotheses were as follows:

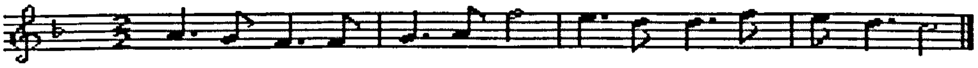
1. Rhythm recall performance would be optimal for sequences with meaningful and repetitive rhythmic and pitch structure (original melody).
2. Sequences with both rhythmic and pitch structures reversed, would be the most difficult to recall.

Excerpts were unaccented and were played at a tempo of one crotchet beat per 0.60 seconds. Sequences were heard binaurally through headphones at comfortable listening levels.

Procedure Listeners were split into two main groups and run individually in the presence of the experimenter in a quiet room. The experimental procedure was the same as in Experiment I. Sequences were presented in a fully randomised order and listeners in each group performed on all sequences for that group. Stimuli presented to each group differed slightly. Listeners proceeded at their own pace with pauses as desired. The experiment took thirty minutes on average.

Analysis Data was scored using the ratio grading method described above in Experiment I (pp 107 - 108). An example of stimuli and scored response is shown below in Figure IIb.

Figure IIb Sample stimuli and 70 % accurate scored response



Stimuli durations	12	4	12	4	12	4	16	12	4	12	4	4	12	16
Response durations	40	13	38	10	37	11	51	37	14	39	12	35	12	40
Stimuli ratios	0.33	3	0.33	3	0.33	4	0.75	0.33	3	0.33	1	3	1.33	
Response ratios	0.33	2.92	0.26	3.70	0.29	4.64	0.72	0.37	2.79	0.31	2.91	0.34	3.33	
Scored analysis I	1	1	1	1	1	1	1	1	1	1	0	0	0	
Scored analysis II	1	1	1	1	1	1	1	1	1	1	0	0	0	
Score notes	11 (correct ratios + 1)													
Score crotchet beats	12 (total correct stimuli durations (12+4+12+4+12+4+16+12+4+12+4 = 96) / crotchet beat duration (8))													

Results

A series of Spearmans correlations were performed on the data to examine the relationship between the two different scoring methods. Responses were grouped by stimuli and took scoring method (Score I (stringent): marked up to the first error only or Score II (less stringent): allowing single isolated errors) and accuracy level (80 %, 70 % or 60 %) into account.

Table 8 Method of Scoring correlations

Group 1 participants			Group 2 participants		
Accuracy	Variable	**Score II**	Accuracy	Variable	**Score II**
80 %	**Score I**	0.984	80 %	**Score I**	0.963
70 %	**Score I**	0.980	70 %	**Score I**	0.980
60 %	**Score I**	0.963	60 %	**Score I**	0.953

All p < 0.0001

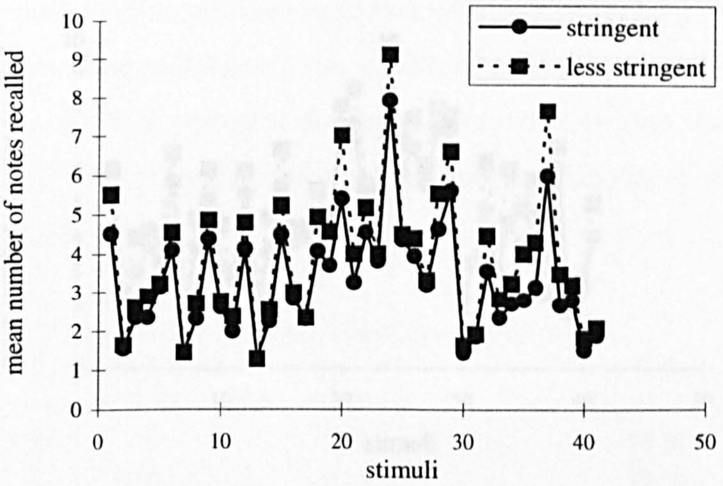


Figure 6 Memory for musical sequences as a function of stimuli at 80% accuracy, group 2

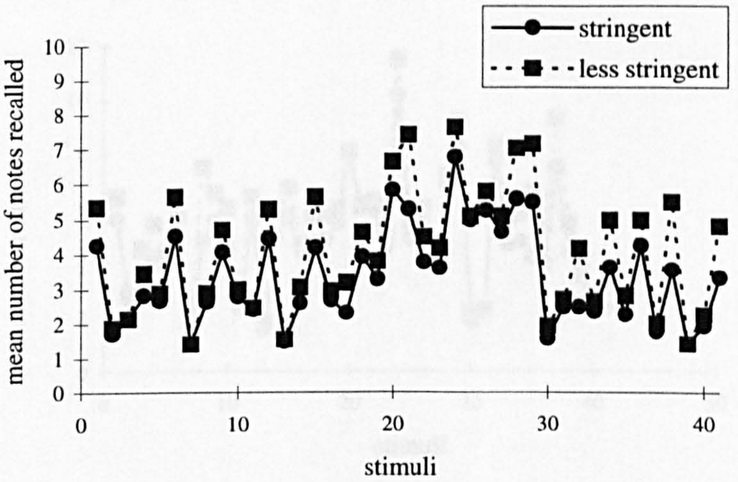


Figure 7 Memory for musical sequences as a function of stimuli at 80% accuracy, group 1

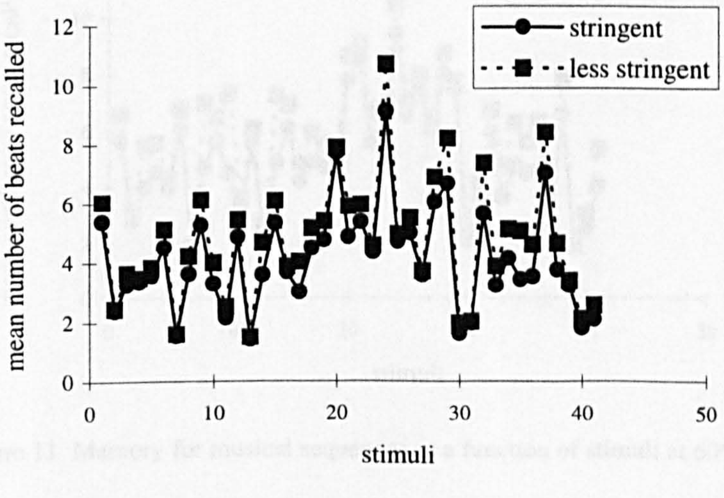


Figure 8 Memory for musical sequences as a function of stimuli at 70 % accuracy, group 2

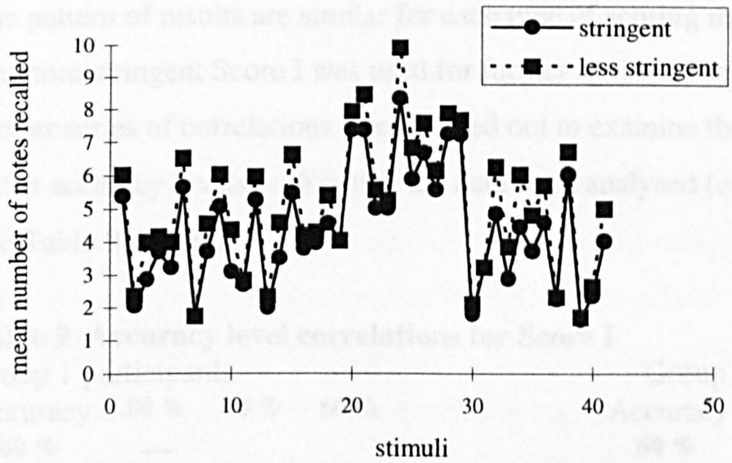


Figure 9 Memory for musical sequences as a function of stimuli at 70% accuracy, group 1

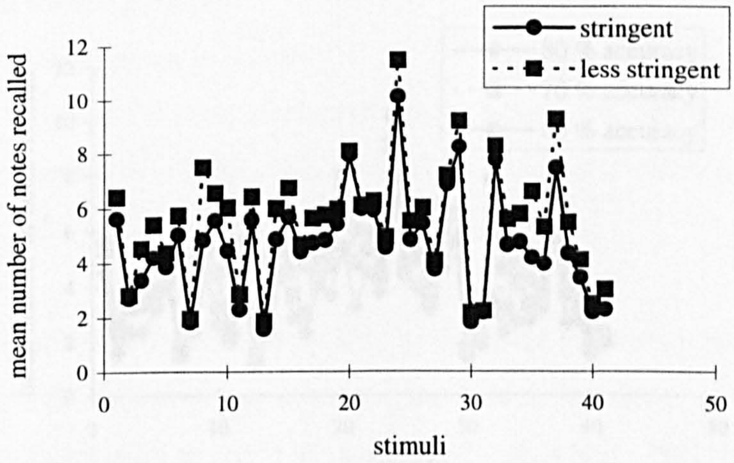


Figure 10 Memory for musical sequences as a function of stimuli at 60 % accuracy, group 2

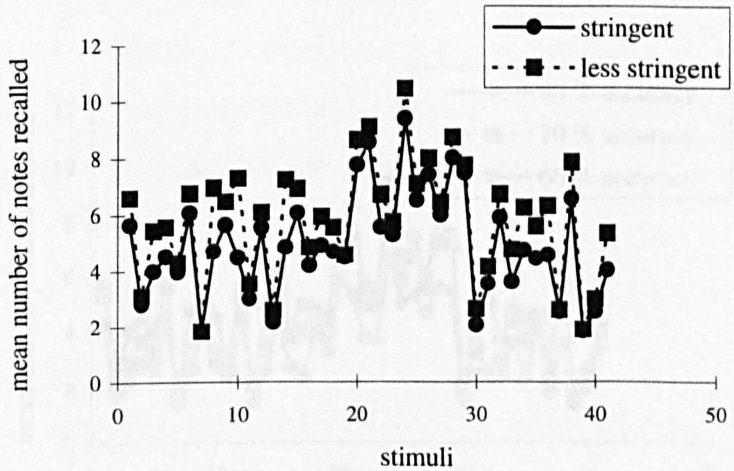


Figure 11 Memory for musical sequences as a function of stimuli at 60% accuracy, group 1

The pattern of results are similar for each type of scoring method. As in Experiment I, the more stringent Score I was used for further statistical investigation of the data. A further series of correlations were carried out to examine the relationship between each of the accuracy levels with which the data were analysed (either 80 %, 70 % or 60 %). See Table 9, below.

Table 9 Accuracy level correlations for Score I

Group 1 participants

Accuracy	**80 %**	**70 %**	**60 %**
80 %	---		
70 %	0.962	---	
60 %	0.896	0.967	---

Group 2 participants

Accuracy	**80 %**	**70 %**	**60 %**
80 %	---		
70 %	0.961	---	
60 %	0.924	0.981	---

All $p < 0.0001$

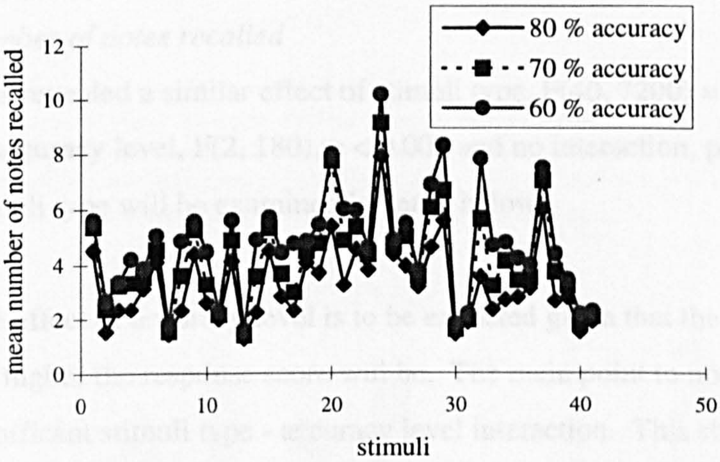


Figure 12 Memory for musical sequences as a function of stimuli type, group 2

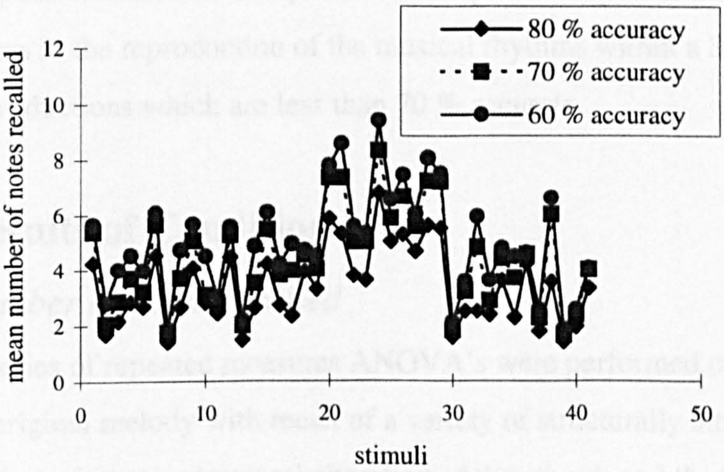


Figure 13 Memory for musical sequences as a function of stimuli type, group 1

Again, the pattern of responses for each level of accuracy are highly positively correlated. As has been noted, the interest lies in examining the relative differences in recall of different stimuli. The important point in choosing one accuracy level for the statistical analysis is that the accuracy level should not determine the statistical effect. All results should be the same regardless of which accuracy level is chosen to score the data. A preliminary repeated measures ANOVA was performed on the recall responses, with stimuli type as the within-subjects factor and accuracy level as the between-subjects factor to determine whether this was the case.

Number of beats recalled

There was a significant effect of stimuli type, $F(40, 7200) = 58.93$, $p < 0.001$, an effect of accuracy level, $F(2, 180) = 14.53$, $p < 0.001$ and no stimuli type - accuracy level interaction, $p > 0.70$.

Number of notes recalled

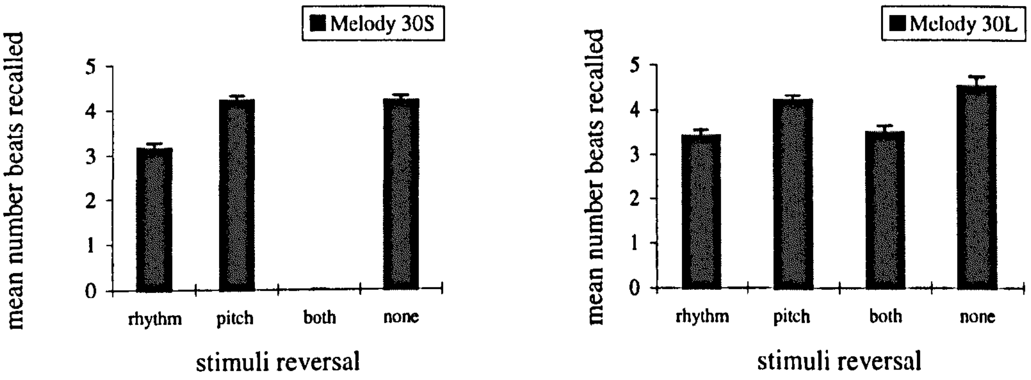
This revealed a similar effect of stimuli type, $F(40, 7200) = 50.26$, $p < 0.001$, an effect of accuracy level, $F(2, 180)$, $p < 0.001$ and no interaction, $p > 0.10$. The effect of stimuli type will be examined in detail below.

The effect of accuracy level is to be expected given that the more lenient a given level is the higher the response score will be. The main point to note is that there is no significant stimuli type - accuracy level interaction. This shows that any statistical effects in the data will be consistent for all accuracy levels. In other words, accuracy level does not determine the results. The accuracy level of 70 % was used for the purpose of statistical comparison in the present experiment. This levels allows for errors in the reproduction of the musical rhythms within a 30 % range and rejects reproductions which are less than 70 % accurate.

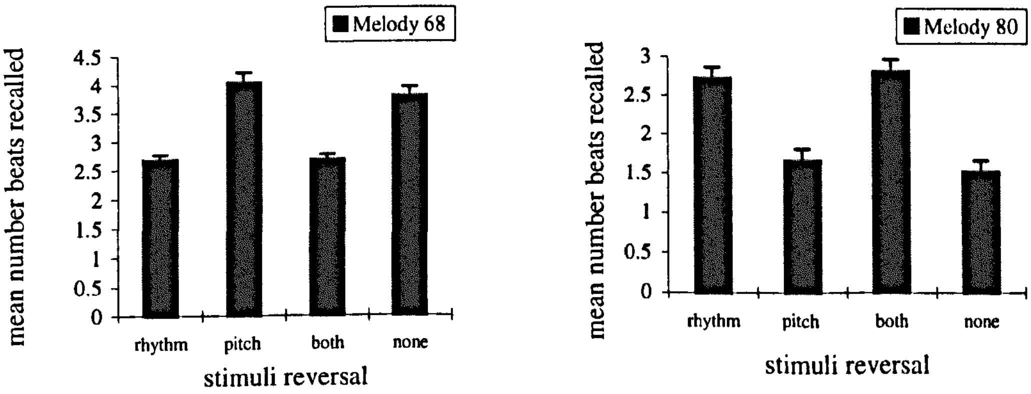
Results of Condition I

Number of beats recalled

A series of repeated measures ANOVA's were performed on the data to compare recall of original melody with recall of a variety of structurally altered sequences. The within-subjects factor is structural alteration of the stimuli and the between-subjects factor is subject group, either 1 or 2. In cases where one group only performed with a given stimulus, the between-subjects factor is subjects. The results are tabulated for each condition below (Tables 10a - 10f, Appendix A, pp 140 - 141). Graphs displaying the mean results for each melody are displayed in Figures 14a - 18, below. Missing bars in the graphs show that the particular missing condition was not run for that melody.



Figures 14a & 14b Mean recall performance as a function of structural alteration of 30S and 30L. The y-axis displays the mean number of beats recalled under each of the conditions. The x-axis displays the structural alterations made to the melodic stimuli: *rhythm* refers to the rhythm only reversed condition; *pitch* to the pitch only reversed condition; *both* to the condition which reversed both pitch and rhythmic structure at the same time and *none* refers to the original melody.



Figures 15 & 16 Mean recall performance as a function of structural alteration of melodies 68 and 80. As detailed above, the y-axis shows the mean number of beats recalled and the x-axis displays the type of stimuli reversal.

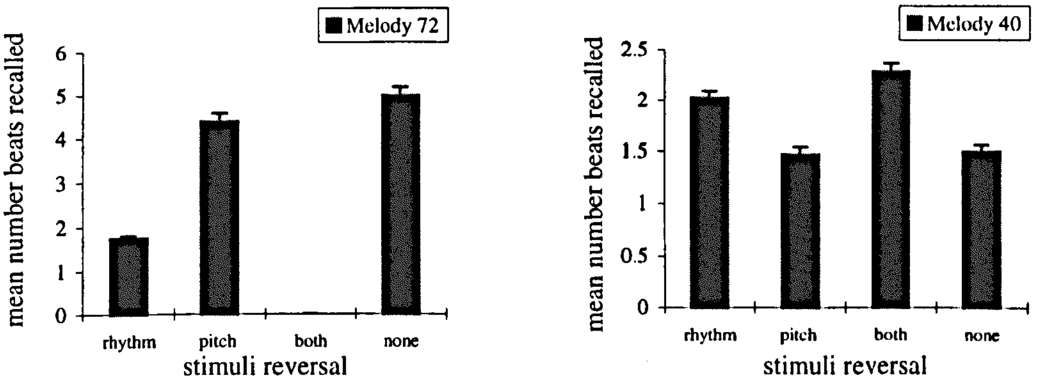


Figure 17 & 18 Mean recall performance as a function of structural alteration of melodies 72 and 40. As detailed above, the y-axis shows the mean number of beats recalled whilst the x-axis displays the type of structural reversal.

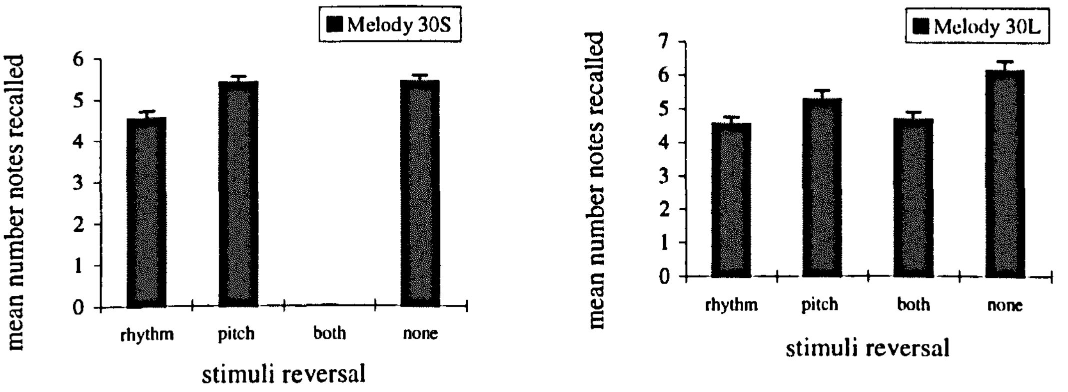
Summary of beat analysis

The beat analysis revealed the following results. Conditions of reversed rhythmic structure (i.e. both reversed and rhythm reversed) yielded significantly poorer recall results than recall of original melodies, 30S, 30L, 68 & 72. For the remaining 2 melody excerpts, 80 and 40, conditions of reversed rhythmic structure were easier to recall than conditions of preserved rhythmic structure.

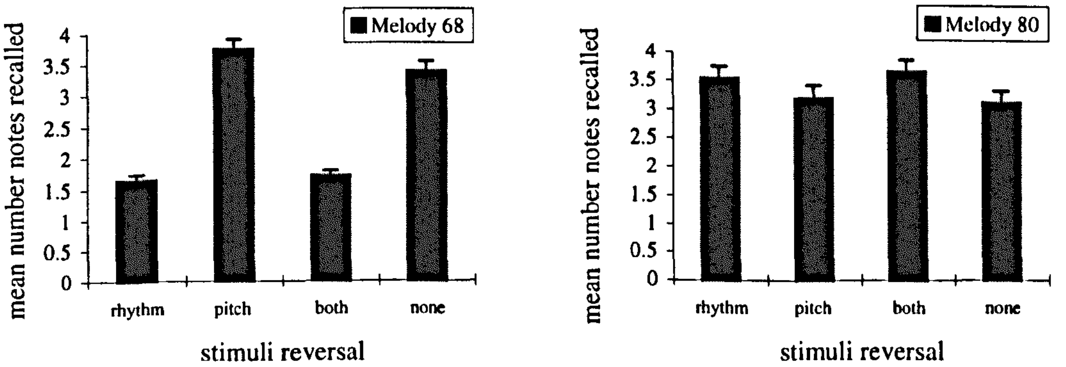
For all 6 melody excerpts, there was no significant difference between recall of original melody and recall of pitch reversed melody. Pitch reversal did not affect recall performance in relation to recall performance of the original melody. There were significant differences between conditions of reversed rhythmic structure and conditions of reversed pitch structure alone for melodies 30S, 72 & 68 with poorer recall for conditions of reversed rhythmic structure. The differences between pitch only reversals and conditions of reversed rhythmic structure were also significant for melodies 80 and 40. However, in these cases the pitch only conditions were significantly *harder* to recall than conditions with reversed rhythmic structure. There were no significant differences between recall of both reversed and rhythm reversed for any of the melodies.

Number of notes recalled

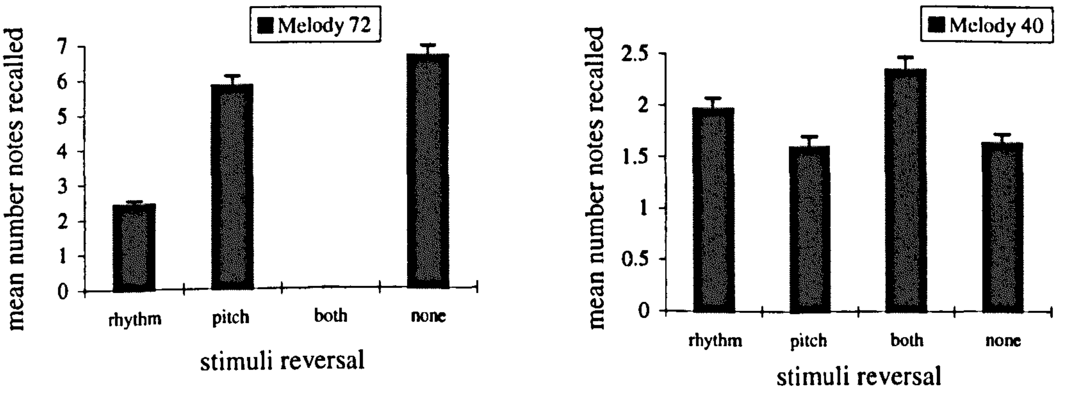
The results of a series of repeated measures and independent ANOVA's are shown below in Tables 11a - 11f, Appendix B, pp 141 - 142. As above in the beat analysis, the within-subjects factor is structural alteration of the stimuli and the between-subjects factor is subject group, either 1 or 2. In cases where one group only performed with a given stimulus, the between-subjects factor is subjects. Mean recall scores for each melody are displayed below in Figures 19a - 23.



Figures 19a & 19b Mean recall performance as a function of structural alteration of 30S and 30L. The y-axis shows the mean number of notes recalled and the x-axis displays the type of stimuli reversal.



Figures 20 & 21 Mean recall performance as a function of structural alteration of melodies 68 and 80. The y-axis shows the mean number of notes recalled and the x-axis displays the type of stimuli reversal.



Figures 22 & 23 Mean recall performance as a function of structural alteration of melodies 72 and 40. The y-axis shows the mean number of notes recalled and the x-axis displays the type of stimuli reversal.

Summary of note analysis

The note analysis revealed a similar pattern of results to the beat analysis. Conditions with reversed rhythmic structure were significantly harder to recall than original melodies 30S, 30L, 68 & 72. The both reversed condition was significantly easier to recall than original melody 40, while there was no significant difference between recall of the rhythm alone reversed and original melody 40. There were no significant differences in number of notes recalled for melody 80 for any of the structural comparisons. Again, there were no significant differences between pitch reversals and original melodies for all of the excerpts. Pitch reversed conditions were significantly easier than conditions with reversed rhythmic structure in melodies 30S, 68 & 72. The differences between pitch reversed and rhythm reversed conditions were not significant for the remaining melodies. Again, there were no significant differences between both reversed and rhythm only reversed conditions for any of the melodies.

Melodies 72 and 30L were the easiest to recall with recall of over 5 beats worth of the rhythmic structure remembered. Melodies 30S and 68 were well remembered with performances in the region of 4 beats. The remaining melodies were harder to recall, with an average of less than 3 beats worth of information remembered.

Condition II Is rhythm recall of melody different to recall of rhythm alone?

Listeners also heard the rhythms of four of the melodies alone without pitch patterns (melodies 72, 68, 30L & 80, see Figure IIc, Appendix 2, page 139) and performed the rhythm recall task on these rhythm sequences. The responses to the rhythm alone sequences were compared with each other to investigate which rhythmic structure facilitated recall. In addition to this, response to the rhythm alone sequences were compared with responses to the original melody to examine the effect of isolating the rhythmic structure on memory for the rhythm of these sequences. The results of these analyses are shown below.

1. Rhythmic structure: which rhythm facilitates recall performance?

Number of beats recalled

Recall of rhythm 72 was significantly higher than recall of rhythms: 30L with $F(1, 60) = 6.90$, $p < 0.05$; 68 $F(1, 60) = 5.52$, $p < 0.05$ and 80 with $F(1, 60) = 65.72$,

$p < 0.001$. There were no significant group effects and no interactions. Means are displayed in Figure 23a below. There was no significant difference between rhythms 30L and 68, $p > 0.70$. Recall of rhythm 30L was significantly better than rhythm 80, $F(1, 60) = 50.06$, $p < 0.001$. Rhythm 68 was significantly easier than rhythm 80, $F(1, 60) = 45.68$, $p < 0.001$.

Number of notes recalled

Again, recall of rhythm 72 was significantly higher than recall of all other rhythms: 30L, $F(1, 60) = 6.66$, $p < 0.05$; 68 with $F(1, 60) = 28.94$, $p < 0.001$ and 80 $F(1, 60) = 35.73$, $p < 0.001$. There were no group effects, $p > 0.40$ and no interactions, $p > 0.10$. Means are shown in Figure 23b, below. There were significant differences between rhythms 30L and 68, $F(1, 60) = 14.37$, $p < 0.001$ and rhythms 30L and 80, $F(1, 60) = 22.45$, $p < 0.001$. Recall of rhythm 30L was significantly higher than melodies 68 and 80. There were no significant differences between melodies 68 and 80, $p > 0.10$.

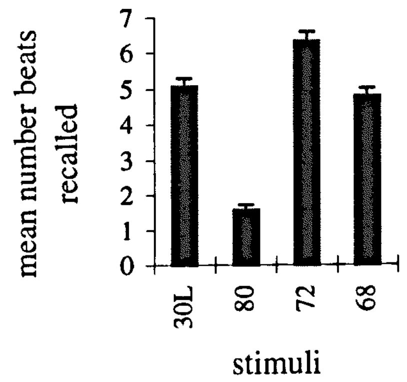


Figure 23a Mean number of beats recalled as a function of rhythmic structure.

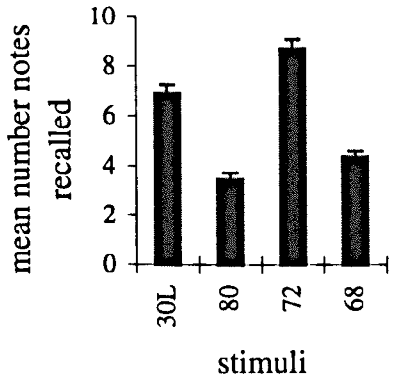


Figure 23b Mean number of notes recalled as a function of rhythmic structure.

In summary, the beat analysis showed that recall was most successful for rhythm 72, performance was less high for rhythms 30L and 68 and recall was poorest for rhythm 80. The note analysis also revealed that the number of notes recalled was highest for rhythm 72. Listeners recalled the second highest number of notes from rhythm 30L and recalled the lowest number of notes from rhythms 68 and 80.

2. Is rhythm recall of melodies different to recall of rhythmic structure alone?

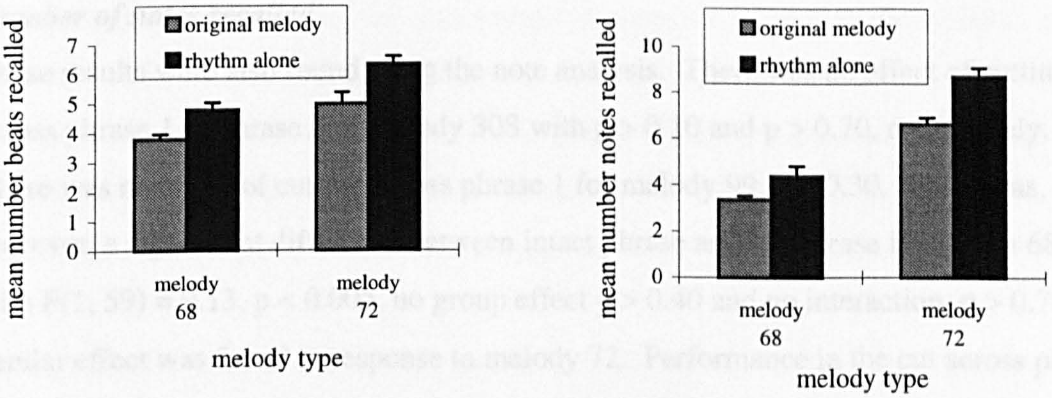
Number of beats recalled

There was no significant difference between recall of melodies 30L and 80 and recall of the rhythmic structure of these melodies alone without pitch patterning with $p > 0.10$.

Rhythm recall performance for rhythmic structure alone was similar to recall of original melody for these two melodies. However, presenting the rhythmic structure of melody 68 alone caused a significant increase in recall, F(1, 59) = 4.52, $p < 0.05$, with no group effect $p > 0.30$, and no interaction $p > 0.50$. Recall was also facilitated by presenting the rhythmic structure alone for melody 72, F(1, 60) = 6.80, $p < 0.05$, with no group effect $p > 0.10$ and no interaction $p > 0.80$. Means are displayed in Figure 24a below.

Number of notes recalled

The same analyses was performed on the number of notes recalled with the following results. There was no significant difference between recall of rhythm alone and recall of melodies 30L and 80, $p > 0.09$. Presenting the rhythmic structure of melody 68 alone facilitated recall of the sequence, F(1, 59) = 4.52, $p < 0.05$, with no effect of group $p > 0.30$ and no interaction $p > 0.50$. Presenting the rhythmic structure alone without pitch patterning had a significant effect on recall for melody 72, F(1, 60) = 7.51, $p < 0.01$, with no group effect $p > 0.80$ and no interaction $p > 0.20$. Recall of rhythm alone was better than recall of the original melody. Means are displayed in Figure 24b, below.



Figures 24a & 24b The beneficial effect of isolating rhythmic structure on memory for the rhythms of musical sequences.
The y-axes show the mean number of beats and notes recalled, respectively. The x-axes display melody type (melody 68 and melody 72). The grey data bars show mean recall of the original melody (both pitch and rhythm patterns intact) whilst the black data bars show mean recall of the rhythmic structure of the melody (rhythm pattern alone on one note).

Condition III Recall of complete phrase vs incomplete phrase sequences.

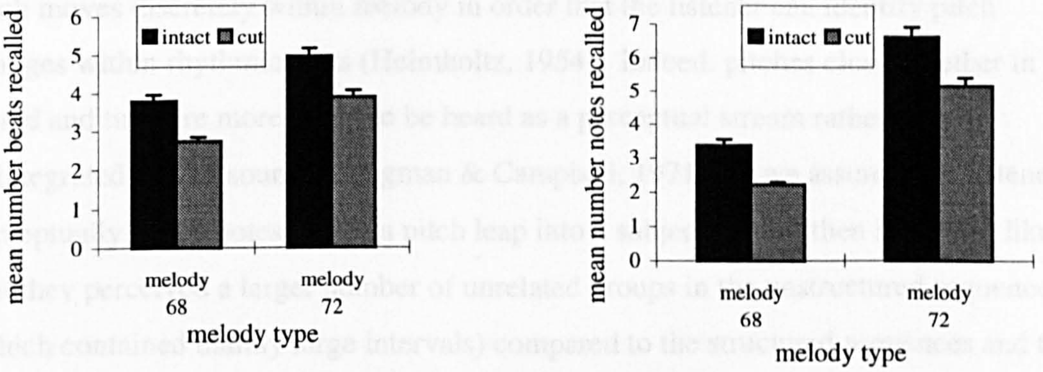
Conditions were run to examine whether presenting stimuli which cut across phrase boundaries (*cut*) were harder to recall than corresponding stimuli from the same melody, identical in length (same number of notes and crotchet beats) which preserved the phrase (*intact*). The four excerpts used as experimental stimuli are displayed in Figure IId, Appendix 3, page 139. As discussed in chapter 2, listeners tended to process information for these melodies in phrasal units. For the present experiment it was hypothesised that excerpts which preserved the phrase as intact would be easier to recall than excerpts which cut across a given phrase.

Number of beats recalled

There was no effect of cutting across phrase 1 in melody 30S on recall performance, $p > 0.20$. This result was also found in phrase 2 of melody 30S, $p > 0.20$ and in phrase 1 of melody 99, $p > 0.40$. There was no benefit to recall of hearing these phrases intact in comparison to cut. There was, however, a significant effect of cutting across the phrase of melody 68, $F(1, 60) = 9.13$, $p < 0.005$, with no group effect $p > 0.30$ and no interaction $p > 0.70$. There was also a significant effect of cutting across the phrase in melody 72, $F(1, 59) = 8.98$, $p < 0.005$, no group effect $p > 0.60$ and no interaction $p > 0.10$. Significantly less of the cut melody was recalled in comparison to the intact melody. Means are shown in Figure 25a, below.

Number of notes recalled

These results were also found using the note analysis. There was no effect of cutting across phrase 1 or phrase 2 of melody 30S with $p > 0.10$ and $p > 0.70$, respectively. There was no effect of cutting across phrase 1 for melody 99, $p > 0.30$. There was, however, a significant difference between intact phrase and cut phrase in melody 68, with $F(1, 59) = 9.13$, $p < 0.005$, no group effect $p > 0.40$ and no interaction, $p > 0.70$. A similar effect was found in response to melody 72. Performance in the cut across phrase condition was significantly different to the intact phrase condition, $F(1, 59) = 7.98$, $p < 0.01$, no group effect $p > 0.20$ and no interaction, $p > 0.10$. A greater number of notes were recalled from the intact phrase than the cut phrase (Means, see Figure 25b, below).



Figures 25a & 25b The beneficial effect of intact phrase structure on memory for the rhythms of musical sequences.
The y-axes of 25a & 25b show the mean number of beats and notes recalled, respectively. The x-axes display melody type (melody 68 and melody 72). The grey data bars show mean recall of the cut melody whilst the black data bars show mean recall of the rhythmic structure of the intact melody.

Discussion

The effect of pitch structure on rhythm recall

Experiment II showed that pitch reversal did not effect recall of any of the melodies. This is in contrast to Experiment I which found a significant effect of pitch interval structure on rhythm recall. There are a number of reasons for these results. Pitch reversal of the Gaelic melodies resulted in ascending and descending intervals becoming descending and ascending intervals, respectively, and also occurring in reversed order. Although the reversed pitch structure may not have been congruent with the rhythmic structure in the way that the original melody was, the number of pitch intervals and the size of these intervals was the same as the original melody. This could explain the similarity in performance of original melody and pitch reversed sequence.

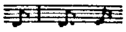
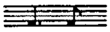
There are differences between experimental stimuli used in the two experiments. Consider the rules used to generate the pitch patterns in Experiment I (described above on page 105 - 106). The structured sequence consisted of tonal runs with pitch intervals of 1 - 4 semitones whereas the unstructured sequence proceeded in atonal pitch intervals which were predominantly larger than 4 semitones. The Gaelic melodies contained pitch intervals of 0, 1, 2, 3 semitones with a probability of 0.267, 0.067, 0.47 and 0.16, respectively. Thus there was a very high probability of the Gaelic melodies containing small pitch intervals (0 - 3 semitones, probability = 0.964).

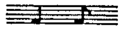
Pitch moves discretely within melody in order that the listener can identify pitch changes within rhythmic units (Helmholtz, 1954). Indeed, pitches close together in sound and time are more likely to be heard as a perceptual stream rather than an unintegrated mix of sounds (Bregman & Campbell, 1971). If we assume that listeners perceptually group notes before a pitch leap into a subjective unit then it is more likely that they perceived a larger number of unrelated groups in the unstructured sequences (which contained mainly large intervals) compared to the structured sequences and the Gaelic melodies (which contained mainly small intervals). This may account for the surprising finding that recall of the unstructured sequences was poorer than the structured sequences. This could also explain the similarity in recall of pitch reversed and original Gaelic melodies which contained the same size of small pitch intervals (0 - 3 semitones) but in reversed order.

The disruptive effect of the unstructured sequences on rhythm recall in Experiment I is most likely due to the fact that the tones did not belong to any specific key and the pitch intervals were predominantly larger than 4 semitones. Studies have shown that melodies generally consist of small pitch intervals, such as minor and major seconds, between adjacent notes and that intervals larger than 4 or 5 semitones are relatively uncommon (Vos & Troost, 1989; Shepard, 1982; Dowling & Harwood, 1986). The normal grouping process which listeners engage in during the course of their music experience would have been disrupted by the unstructured pitch patterns. The pitch patterns were not ordered in any systematic manner and the unstructured sequences did not conform to the structural patterns with which listeners are familiar. It is possible that music which does not conform to the structural patterns with which a listener is used to hearing is more difficult to perceive, process and represent in memory because listeners are unfamiliar with the melodies and cannot group and recursively link the rhythmic groups of information. Recall of the Gaelic melody rhythms used in the present experiment might also have been disrupted if they had been matched with the pitch interval structure of the unstructured sequences.

The results of Experiment I suggested that STM for the rhythms of auditory sequences is facilitated when the pitch pattern is structured with tonal runs containing small pitch intervals. The implications of this are that pitch and rhythm factors are interdependent in the perception of musical sequences. As has been noted, the pitch and rhythm patterns were constructed separately by simple rules which may have resulted in melodies containing incompatible temporal and pitch features. Rhythm and pitch patterns which are compatible will be highly coherent and should facilitate encoding and remembering (Boltz, 1989).

Another factor to consider is that the repetitive nature of the rhythm patterns in the Gaelic melodies may make them easier to remember than the computer generated rhythms used in Experiment I. It is interesting to note that the average recall of computer generated 3 - 7 beat sequences was generally a function of increasing sequence length (See Figures 1a - 3b, pages 112 - 113). As sequence length increased, recall decreased. This was not the case for rhythm recall of the Gaelic melodies which did not decrease with increasing sequence length. The recall results highlight the role of repetition as an important structural element within real melodies. Simple patterns within melodies may actively enhance the formation of larger perceptual units. When successive items are unrelated, memory decreases as number of items increases (Krumhansl, 1990). This was result the of Experiment I: an average of 3 beats or 4 notes was recalled from structured computer generated sequences of 4 beats in length (at 70 % accuracy level) with an overall decrease in recall with increasing sequence length. The average recall of the Gaelic melody rhythms in Experiment II, which contained well-defined repeating units, was in the higher region of 4 - 5 beats or 5 - 7 notes. The relative ease of recall of the Gaelic rhythms which contained repetitive units may have reduced any potentially detrimental effect of mismatched pitch structure. These explanations could also account for the lack of pitch structure effects found in Experiment II.

The main results of Experiment II suggest that sequences containing repetitive rhythmic units are easier to recall than sequences without such rhythm patterns. It was shown that most of the excerpts containing reversed rhythmic structure (conditions both & rhythm reversed) were harder to recall than excerpts which preserved the rhythmic structure (original melody & pitch reversed). This highlights the importance of rhythmic structure in facilitating memory for rhythmic information. The common feature of melodies whose original rhythmic structure facilitated recall is the repeating units from the start of each melody. Melody 72 consists of a series of the repeating unit , whereas the reversed melody does not contain a repeating unit. Melody 68 contains a series of the repeating unit , while the reversed excerpt contains no simple repeating pattern at the start of the excerpt. The similarity between recall of reversed pitch excerpts (which preserve the rhythmic structure) and recall of original melody excerpts supports the theory that it is the absence of repeating rhythm units which causes the reduction in recall. The finding that recall of pitch reversed sequences is significantly better than recall of sequences which reversed the rhythmic structure in most of the melodies is additional support for this interpretation. The unexpected finding is that reversing the rhythmic structure enhanced recall for melodies 80 and 40. This result can be accounted for by a closer inspection of the musical structure. Rhythm reversal does not necessarily result in the break-down of rhythmic structure, indeed reversal for melody 80 resulted in an easier rhythmic pattern at the start of the excerpt. Melody 80 began with the difficult 1:1.5 ratio  which many listeners may have misperceived and erroneously recalled. Whereas in the reversed sequence the melody began with a 3: 1 ratio  which is easier to perceive and thus produce (Gabrielsson, 1974; Povel, 1981).

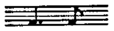
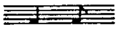
So what were the features of the melodies which were most successfully recalled by listeners? The recall responses can be considered in relation to pitch and temporal patterns in the specific melodies. Melodies 72 & 30L resulted in the highest recall performance. An average of around 7 notes were recalled for these melodies which is in line with previous studies which have examined memory limitations (Miller, 1956). The common characteristic of these melodies is the series of repeating  units which implies that these repeating units can enhance recall. It is interesting to note that the average number of notes recalled from 30L and 72 corresponds to the number of notes in the repeating pattern before the rhythm changes to consecutive quavers. The change in repeating pattern may have disrupted listeners processing of the sequence.

In addition to this, the number of different note durations in the melodies can effect rhythm recall. Melody 68 contains three different durations (minim, dotted crotchet and quaver) whilst melodies 72 and 30L contain only two different durations (quaver and crotchet). Previous research has shown that rhythm reproduction of sequences containing three durations (e.g. crotchet, quaver, semi-quaver) is harder than rhythm reproduction of two note durations (e.g. quaver, semi-quaver) (Drake, 1993). The present results support this finding in that melody 68 is poorly recalled in comparison to melodies 72 and 30L.

Does pitch pattern enhance rhythm recall of real melodies?

Rhythm recall of rhythm sequences (*without* pitch information) was compared with the corresponding melody (the same rhythm pattern *with* pitch information) and yielded mixed results. The rhythm alone sequences were better recalled than two of the corresponding melodies. A greater number of beats and notes were recalled from melody 72 on the rhythm alone condition. Thus the pitch sequence did not facilitate rhythm recall. Listeners also recalled more rhythm information from the rhythm alone condition than from original melody 68. The facilitatory effect of hearing the rhythm alone was not, however, replicated in the remaining two melodies in which recall of rhythm alone and of melody did not differ. The same amount of rhythmic information was recalled from rhythm 30L as from melody 30L and from rhythm 80 as from melody 80. There is no neat explanation for the facilitatory effect of the rhythmic structure alone in comparison to melodies 72 and 68 and not for melodies 30L and 80. One difference between these melodies is that 30L and 80 contain repeating pitches at the start of the melody so the rhythm alone and melody are similar. This is not the case for melodies 72 and 68 which may account for the contrasting results of rhythm alone versus melody on rhythm recall.

Recall of intact and cut phrases

The results of recall of sequences which cut across phrases and recall of sequences which preserved the phrase were mixed. There was no effect of cutting across phrase boundaries for melodies 30S and 99 but there was a significant facilitatory effect of preserving the phrase structure of melodies 68 and 72. These results can be explained if we consider the structural properties of the melodies. The stimuli which cut across phrases in melodies 68 and 72 also resulted in large pitch leaps which occurred earlier than in the excerpts of the complete phrase. If, as suggested above, listeners perceptually group notes up until a pitch leap into a memorable unit, then the incomplete phrase contains a shorter memorable unit than the complete phrase. In addition to this, there are fewer repeating units in the cut sequences. Melody 68 contains 3 lots of  before a minim, whereas 68 cut contains only 1 lot of  before a minim occurrence. Similarly, melody 72 contains 6 lots of repeating  units before the pattern changes to consecutive quavers, whereas 72 cut contains only 4 lots of repeating  units before the pattern changes to consecutive quavers. Thus reduced recall in cut melodies could result from smaller numbers of repeating units before a change in the rhythm pattern and from large unexpected pitch leaps early on in the sequence which result in a shorter memorable group.

Summary

The present findings show that, despite employing the same rhythm recall task, different stimuli can produce contradictory results. In line with previous experiments which demonstrate that pitch and rhythm features are interrelated in the perception of simple melodies, the results of Experiment I (tonal with small interval versus atonal with large interval sequences) show that rhythm recall was affected by the pitch interval structure of the sequences. This finding was not, however, replicated in Experiment II (Gaelic melody versus pitch reversed melody). Rhythm recall of Gaelic melodies was unaffected by the reversal of pitch structure. A closer examination of the stimuli utilised in the two experiments can account for these results. The effect of pitch pattern on rhythm recall which was reported in Experiment I may have been due to the fact that the unstructured sequences were characterised by particularly large atonal pitch intervals while the structured sequence contained small tonal pitch intervals. The lack of an effect of pitch structure reversal of the Gaelic melodies can be explained by the similarity in pitch interval size of the original and reversed melody (both sequences contained small pitch intervals of 0 - 3 semitones).

In addition to this the artificial, computer-generated rhythms lacked the meaningful and memorable repetitive rhythm units which were inherent within the Gaelic rhythms so the computer generated rhythms were not as memorable. This, combined with the particularly extreme pitch structure of the unstructured sequences, also explains the reduced recall obtained in Experiment I. In other words, the rhythms without repeating rhythmic units are hard to recall and when combined with large atonal pitch leaps, which further distract the listener from the already demanding rhythm recall task, recall is disrupted. This is an interesting finding, given that listeners' attention is focused primarily on the rhythmic structure when performing the present rhythm recall task.

As noted above, the lack of a pitch reversal effect in Experiment II may result from the fact that the pitch reversals resulted in pitch patterns which contained the same pitch intervals as those prevalent in the original melody. In addition to this, the rhythmic structure of the Gaelic melodies is repetitive and is likely to be easier to process and remember than the non-repetitive computer generated rhythms of Experiment I. It may be that the pitch patterns of sequences which contain memorable repetitive rhythm units are less distracting than the pitches of sequences which do not contain such memorable rhythmic units. It would be interesting to examine whether recall of the Gaelic melody rhythms which contain repetitive rhythmic units would be poorer if accompanied by the same pitch interval structure of the unstructured sequences.

Experiments I and II generated contrasting findings for how pitch and rhythm are processed by the listener. However, a closer inspection of the stimuli shows that the contrasting results were obtained as a result of the different and in some cases extreme (atonal sequences of Experiment I) types of experimental stimuli utilised in the experiments. Neither experiment warrants the conclusion that pitch and temporal factors are/are not independent in the perception of melodies. The present results imply that particular task demands must be taken into account in conjunction with the particular stimuli utilised in the experiment. Care must be taken to interpret appropriately the results of experiments which manipulate pitch and rhythm features in the musical stimulus.

The present findings showed that listeners are best at recalling well-structured rhythms with simple pitch patterns. The results suggest that processing of and memory for simple melodies are facilitated by the use of simple repetitive units in the melody.

Appendices for chapter 3

Appendix 1

Figure IIa Stimuli used in rhythm recall Experiment II: Condition I

Melody 30S original

Melody 30S Pitch reversed

Melody 30S Rhythm reversed

Melody 30L original

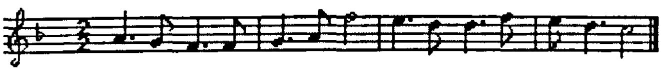
Melody 30L Both reversed

Melody 30L Pitch reversed

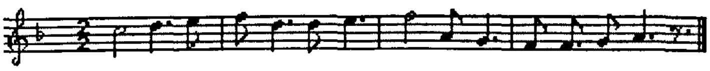
Melody 30L Rhythm reversed

Melody 30L Rhythm alone

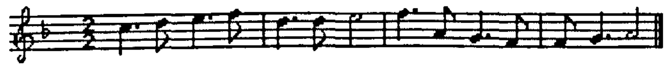
Melody 68 original



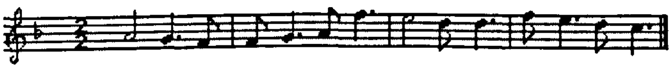
Melody 68 Both reversed



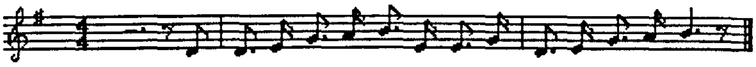
Melody 68 Pitch reversed



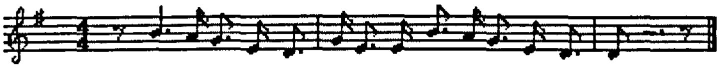
Melody 68 Rhythm reversed

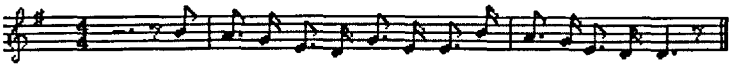


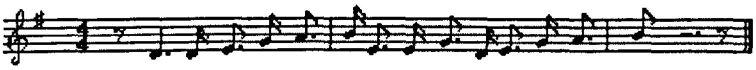
Melody 80 original



Melody 80 Both reversed

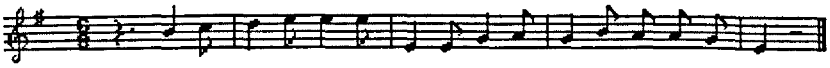


Melody 80 Pitch reversed

Melody 80 Rhythm reversed



Melody 72 original

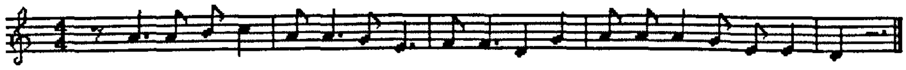
Melody 72 Pitch reversed



Melody 72 Rhythm reversed

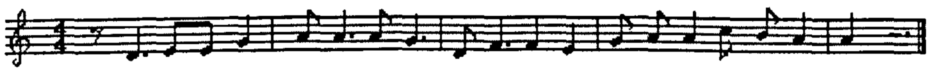
Melody 40 Original

Melody 40 Both reversed



Melody 40 Pitch reversed

Melody 40 Rhythm reversed

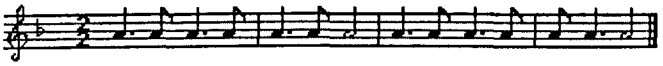


Appendix 2

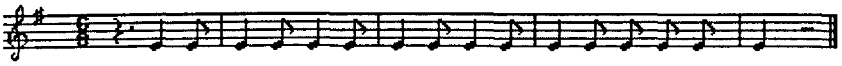
Figure IIc Rhythms of melodies used in rhythm recall Experiment II: Condition II

Rhythm 30L

Rhythm 68



Rhythm 72



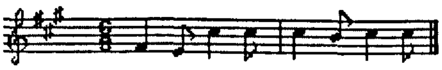
Rhythm 80

Appendix 3

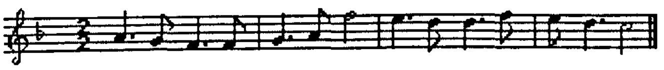
Figure IId Melodies used in the recall experiment: Condition III

Melody 30S original

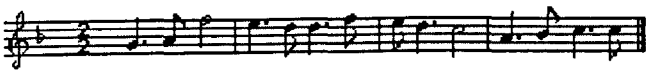
Melody 30S Cut



Melody 68 original



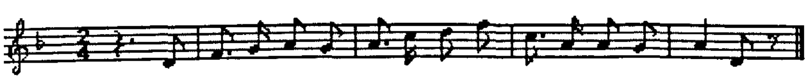
Melody 68 Cut



Melody 72 original

Melody 72 Cut

Melody 99 original



Melody 99 Cut

Appendix A Anova Tables for Experiment II Condition 1 Analysis by beats

Table 10a Original melody vs Both reversed

Melody	Stimuli effect	Group effect	interaction
68	$F(1, 60) = 6.91, p < 0.05$	$p > 0.60$	$p > 0.20$
80	$F(1, 30) = 14.96, p < 0.001$	---	---
40	$F(1, 30) = 21.21, p < 0.001$	---	---
30L	$F(1, 60) = 3.79, p < 0.05$	$p > 0.06$	$p > 0.10$

Table 10b Original melody vs Rhythm alone reversed

Melody	Stimuli effect	Group effect	interaction
68	$F(1, 60) = 10.47, p < 0.005$	$p > 0.90$	$p > 0.40$
80	$F(1, 30) = 8.58, p < 0.01$	---	---
40	$F(1, 30) = 8.79, p < 0.005$*	---	---
30L	$F(1, 60) = 6.23, p < 0.05$	$p < 0.01$	$p > 0.60$
30S	$F(1, 60) = 17.15, p < 0.001$	$p > 0.90$	$p > 0.90$
72	$F(1, 60) = 66.34, p < 0.001$	$p > 0.10$	$p > 0.50$

*** independent anova**

Table 10c Original melody vs Pitch alone reversed

Melody	Stimuli effect	Group effect	interaction
68	$p > 0.50$	---	---
80	$p > 0.40$	---	---
40	$p > 0.80$*	---	---
30L	$p > 0.20$	---	---
30S	$p > 0.80$	---	---
72	$p > 0.10$	---	---

*** independent anova**

Table 10d Both reversed vs Rhythm alone reversed

Melody	Stimuli effect	Group effect	interaction
68	$p > 0.60$	---	---
80	$p > 0.70$	---	---
40	$p > 0.20$*	---	---
30L	$p > 0.80$	---	---

*** independent anova**

Table 10e Both reversed vs Pitch alone reversed

Melody	Stimuli effect	Group effect	interaction
68	$F(1, 60) = 10.42, p < 0.005$	$p > 0.30$	$p > 0.50$
80	$F(1, 60) = 21.30, p < 0.001$	$p > 0.90$	$p > 0.80$
40	$F(1, 60) = 14.86, p < 0.001$*	---	---
30L	$p > 0.20$	---	---

independent anova

Table 10f Rhythm alone reversed vs Pitch alone reversed

Melody	Stimuli effect	Group effect	interaction
68	$F(1, 60) = 14.32, p < 0.001$	$p > 0.70$	$p > 0.80$
80	$F(1, 60) = 13.02, p < 0.001$	$p > 0.80$	$p > 0.90$
40	$F(1, 30) = 7.43, p < 0.05$	---	---
30L	$p > 0.20$	---	---
30S	$F(1, 60) = 58.61, p < 0.001$	$p > 0.07$	$p > 0.20$
72	$F(1, 60) = 13.02, p < 0.001$	$p > 0.80$	$p > 0.80$

Appendix B Anova tables Experiment II Condition 1 Analysis by notes

Table 11a Original melody vs Both reversed

Melody	Stimuli effect	Group effect	interaction
68	$F(1, 59) = 19.23, p < 0.001$	$p > 0.60$	$p > 0.20$
80	$p > 0.20$	---	---
40	$F(1, 30) = 7.76, p < 0.01$	---	---
30L	$F(1, 59) = 5.10, p < 0.05$	$p > 0.06$	$p > 0.10$

Table 11b Original melody vs Rhythm alone reversed

Melody	Stimuli effect	Group effect	interaction
68	$F(1, 59) = 28.04, p < 0.001$	$p > 0.90$	$p > 0.40$
80	$p > 0.40$	---	---
40	$p > 0.20$*	---	---
30L	$F(1, 59) = 5.54, p < 0.05$	$p < 0.01$	$p > 0.60$
30S	$F(1, 59) = 8.40, p < 0.01$	$p > 0.90$	$p > 0.80$
72	$F(1, 60) = 55.56, p < 0.001$	$p > 0.10$	$p > 0.50$

*** independent anova**

Table 11c Original melody vs Pitch alone reversed

Melody	Stimuli effect	Group effect	interaction
68	$p > 0.30$	---	---
80	$p > 0.40$	---	---
40	$p > 0.90$*	---	---
30L	$p > 0.20$	---	---
30S	$p > 0.90$	---	---
72	$p > 0.10$	---	---

*** independent anova**

Table 11d Both reversed vs Rhythm alone reversed

Melody	Stimuli effect	Group effect	interaction
68	$p > 0.70$	---	---
80	$p > 0.80$	---	---
40	$p > 0.20$*	---	---
30L	$p > 0.80$	---	---

*** independent anova**

Table 11e Both reversed vs Pitch alone reversed

Melody	Stimuli effect	Group effect	interaction
68	$F(1, 60) = 28.09$, $p < 0.001$	$p > 0.20$	$p > 0.50$
80	$p > 0.30$	---	---
40	$F(1, 30) = 5.54$, $p < 0.05$*	---	---
30L	$p > 0.30$	---	---

*** independent anova**

Table 11f Rhythm alone reversed vs Pitch alone reversed

Melody	Stimuli effect	Group effect	interaction
68	$F(1, 59) = 35.58$, $p < 0.001$	$p > 0.60$	$p > 0.90$
80	$p > 0.50$	---	---
40	$p > 0.20$	---	---
30L	$p > 0.20$	---	---
30S	$F(1, 60) = 58.61$, $p < 0.001$	$p > 0.07$	$p > 0.20$
72	$F(1, 60) = 13.02$, $p < 0.001$	$p > 0.80$	$p > 0.80$

Appendix C Mean rhythm recall of structured/unstructured musical sequences

Table 1 Mean number of beats and notes recalled from structured/unstructured sequences at 70 % accuracy level.

	Number crotchet beats recalled				Number notes recalled			
	Structured n=31		Unstructured n=31		Structured n=31		Unstructured n=31	
Sequence length (in crotchet beats)	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d
3 beats long	2.77	0.24	2.55	0.34	3.89	0.49	3.48	0.56
4 beats long	2.93	0.57	2.73	0.63	4.60	0.96	4.09	0.99
5 beats long	2.77	0.77	2.55	0.83	4.61	1.49	3.96	1.32
6 beats long	2.69	0.98	2.45	0 .73	4.19	1.57	3.78	1.28
7 beats long	2.47	0.49	2.39	0.77	3.88	0.94	3.70	1.19

Appendix D Mean rhythm recall of Gaelic melodies

Table 1 Mean number of *beats* recalled from Gaelic melodies at 70 % accuracy level.

	original melody		both reversed		rhythm reversed		pitch reversed		rhythm alone	
melody	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.
30S	4.22+	1.72	—	—	3.15+	2.06	4.21+	1.76	---	---
30L	4.52+	3.26	3.49+	2.42	3.42+	2.09	3.97+	2.94	5.08+	3.45
68	3.79+	2.58	2.77+	1.35	2.69+	1.44	4.04+	2.73	4.80~	3.63
80	1.53*	1.55	2.80+	1.60	2.74+	1.55	1.71+	1.77	1.79+	2.13
72	4.97+	3.07	—	—	1.76+	0.94	4.39+	2.89	6.34+	3.94
40	1.50*	0.71	2.29*	0.86	2.03*	0.71	1.47*	0.82	---	---

note --- no condition, + n = 62, ~ n = 61, * n = 31

Table 2 Mean number of *notes* recalled from Gaelic melodies at 70 % accuracy level.

	original melody		both reversed		rhythm reversed		pitch reversed		rhythm alone	
melody	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.
30S	5.39+	2.26	—	—	4.52+	2.75	5.39+	2.29	—	---
30L	6.11+	4.71	4.68+	3.55	4.56+	3.07	5.29+	4.20	6.94+	5.04
68	3.39+	2.47	1.71+	1.36	1.65+	1.36	3.74+	2.63	4.39~	3.46
80	2.87*	3.01	3.66+	3.25	3.55+	3.12	3.19+	3.47	3.48+	4.19
72	6.65+	4.39	—	—	2.44+	1.64	5.82+	4.15	8.73+	5.71
40	1.65*	0.95	2.35*	1.36	1.97*	1.17	1.61*	1.12	---	---

note --- no condition, + n = 62, ~ n = 61, * n = 31

Appendix E Mean recall of intact and cut phrases in four Gaelic melodies

Table 1 Mean number of beat and notes recalled from intact and cut phrases of Gaelic melodies at 70 % accuracy level.

	Number crotchet beats recalled				Number notes recalled			
	Cuts across phrase		Intact phrase		Cuts across phrase		Intact phrase	
Melody	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d
30S	3.97+	1.80	4.52+	3.26	5.08+	2.40	5.39+	2.26
99	2.03*	1.80	2.52*	2.95	3.74*	3.35	4.81*	5.56
68	2.74+	1.58	3.79+	2.58	2.23+	1.22	3.39+	2.47
72	3.89~	2.66	4.97+	3.07	5.18~	3.88	6.65+	4.39

note + n = 62 ~ n = 61, * n = 31

Appendix F Algorithm for producing rhythm sequence (pseudo-code)

```
Prob_Increase_Speed=0.5
Prob_Decrease_Speed=0.5
Note (1) = 1
Position (1) = 0

loop over i
        Note (i) = Note (i - 1)
        if (Note (i) = = 1) and (Position (i) = = 0)
                if Random (0, 1) < Prob_Increase_Speed
                        Note (i) = Note (i) / 2
                        Prob_Increase_Speed = 0.5
                endif
endif

        if (Note (i) = = 0.5) and (Position (i) = = 0)
                if Random (0, 1) < Prob_Increase_Speed
                        Note (i) = Note (i) / 2
                        Prob_Increase_Speed = 0.5
                endif
endif

        if (Note (i) = = 0.25) and (Position (i) = = 0)
                if Random (0, 1) < Prob_Decrease_Speed
                        Note (i) = Note (i) * 2
                        Prob_Decrease_Speed = 0.5
                endif
endif

        if (Note (i) = = 0.5) and (Position (i) = = 0)
                if Random (0, 1) < Prob_Decrease_Speed
                        Note (i) = Note (i) / 2
                        Prob_Decrease_Speed = 0.5
                endif
endif

        if (Position (i) = = 0)
                Prob_Increase_Speed = 1
endif

        if (Position (i) = = 0.25)
                Prob_Decrease_Speed = 2/3 - Prob_Decrease_Speed/3
endif

        Position (i) = fractional_part_of (Position (i-1) + note (i))

endloop
```

CHAPTER 4 Beat Perception

Background

The beat, or regular pulse, is one of the most fundamental features of music (Sloboda, 1985). Beat perception, or the ability to detect a series of regular beats underlying the music, is one of the first perceptual responses to occur in music listening (Bigand, 1993). As such, the study of beat perception should provide us with important insights into the way in which the perception of music develops in the course of listening. The next stage in music listening is the perception of metre, whereby the listener differentiates the series of perceived beats into strong and weak beats (Bigand, 1993; Lerdahl & Jackendoff, 1983). Beat and metre perception are necessary in understanding music (Handel & Lawson, 1983). Indeed, listeners consistently organise a piece of music temporally into metrical units such as beats and bars (Lee, 1985). It is the beat which is the most basic level of regular structure and beat perception occurs before metre perception in music listening. Interest in listeners' metrical interpretations may be due to the relevance of beat perception to our musical understanding. This is not to say that metre perception is not important. Rather the point is that beat perception, which occurs early on in the music listening experience, shapes the listeners perception of the metre in the music. Empirical research has demonstrated that metre is important to the listeners' understanding and memory for melodies (Sloboda & Parker, 1985; Palmer & Krumhansl, 1990). In addition to this, an established metre aids the reproduction of simple musical sequences (Povel, 1981). Indeed it has been demonstrated that the same rhythmic pattern can be performed differently in different metrical contexts (Clarke, 1985).

Tapping experiments have been used widely to study the perception of beat and metre in music (Drake et al, 1997; Handel & Oshinsky, 1981; Handel & Lawson, 1983; Vos et al, 1994; Parncutt, 1994). There are also a number of theoretical attempts to explain how the listener perceives the metre of a given sequence or melody (Longuet-Higgins, 1976; Steedman, 1977; Longuet-Higgins & Lee, 1982; Lee, 1985; 1991; Parncutt, 1994). A number of these studies will be discussed in more detail at a later point.

The beat-tapping task

It is preferable that tasks which are used to examine beat perception empirically, do so in a relatively simple non-obtrusive way, without disruption to the normal listening process. Tapping along with the beat of the music is one of the easiest tasks to perform and is as commonplace as listening (Handel & Lawson, 1983; Handel, 1984). Musical and non-musical participants are able to perform this simple task with no effect of musical training (Handel, 1984). The use of a finger-tapping response is the most natural way to establish the listeners' perception of the beat and underlying metre of the music (Vos et al, 1994). In addition to this, finger-tapping responses are easily produced, repeatable and can be used to define a sequence of inter-response intervals (Wing & Kristofferson, 1973a). It has been demonstrated that the spontaneous tempo of the forefinger is well correlated with that of the hand, leg and arm (Mishima, 1965, as cited in Deutsch, 1982a). Performance in a rhythm reproduction task which employed a finger-tapping response was due primarily to perceptual factors rather than constraints on the operation of the motor system (Summers, Bell & Burns, 1990).

A simple finger tapping task is therefore extremely suitable for an investigation of beat perception. In line with the previous three tasks employed in this thesis (chapters 2 & 3), beat-tapping can be performed by non-musicians. It is useful to employ a beat-tapping task in a perceptual study as it allows participants to respond to their preferred beat level in the music and therefore reflects the beat structure in which the music is perceived (Handel & Oshinsky, 1981; Handel, 1984). What this means is that as long as task instructions are kept simple and unconstrained, listeners should perform the beat tapping task at their preferred rate. As a result, responses should not be biased in any manner and should successfully reflect the perceptual experience of the musical sequences used in the experiment.

Problems of interpretation

The results of tapping tasks have been interpreted in a number of different ways. Many researchers focus on explaining tapping responses in relation to motor response limitations (Wing & Kristofferson, 1973b) and in relation to coding theories which are used to predict the difficulty of the sequence structure (Povel & Collard, 1982). There has been a limited amount of work which addresses the role of perceptual constraints in motor response (Kelso, Holt, Rubin & Kugler, 1981) and there is a genuine need for the role of perceptual factors in tapping response tasks to be examined (Klapp, Hill, Tyler, Martin, Jagacinski & Jones, 1985).

Some of the earlier research to employ finger tapping tasks did so to examine how well listeners could synchronise their tapping response with tones in an experimental sequence. Deutsch (1983) examined the ability of listeners to track rhythm patterns and synchronise their tapping response to patterns of tones. Listeners heard a series of 'blips' in the right ear and another series (at a slightly different pitch) in the left ear. Their task was to perform a synchronous tapping response with the left hand shadowing the rhythm in the left ear and the right hand shadowing the response in the right ear. A variety of different durations of blips were used to create simple rhythm patterns: left and right ear tone durations were related in ratios of 1 against 2, 3, 4 & 5, and a series of polyrhythmic patterns: tones were related in ratios of 2 against 3 & 5, 3 against 4 & 5 and 4 against 5. Results show that simple rhythms are easier to follow, and thus synchronise a tapping response with, than polyrhythms. As noted above, listeners were required to synchronise their tapping response with every event in the stimulus and as a result could not choose which events to tap to. The stimuli used in this experiment were impoverished sequences without pitch features. The results are therefore not directly informative about beat perception in response to real musical stimuli.

Klapp et al (1985) were amongst the first to examine empirically the relationship between perceptual features of the stimuli on tapping performance. They used a perceptual monitoring task to examine how well the listener could shadow simple and complex rhythmic patterns. Listeners heard a pattern and were asked to identify the occurrence of this pattern. They also carried out a number of tapping tasks with sequences employing pitch factors. In line with Deutsch (1983), who used a tapping task, their perceptual task revealed that listeners performed poorly on complex polyrhythmic sequences in comparison to simple rhythmic patterns. In addition to this, they found that one-handed tapping responses were affected in the same way as two-handed tapping responses. This cast doubt on previous interpretations which had used the increased complexity of a two-handed motor response as the major explanation for poor performance on a number of tasks. In addition to this, responses to sequences employing the same rhythm patterns, but different pitches, resulted in significant differences in tapping response. This also supports the inclusion of perceptual, rather than purely motoric, interpretations of responses in beat-tapping tasks.

Others have employed one-handed tapping tasks to examine the listeners rhythmic interpretation of polyrhythmic sequences (Handel & Lawson, 1983; Handel & Oshinsky, 1981). In these cases, instructions were kept simple and unconstrained. Listeners were asked to tap along to the beat of the rhythm, keeping time with the sequence. Results showed that rhythm perception is a function of many interrelated cues such as pitch, duration of events in the sequence, and tempo of the pattern.

The findings of Klapp et al (1985), Handel & Lawson (1983) and Handel & Oshinsky (1981) highlight the benefit of accounting for responses to tapping tasks within a perceptual framework. It should be noted that the stimuli used in these studies were relatively impoverished: pitch features and durations of events in the sequence were varied along the most basic levels.

Empirical beat perception research with music

A limited number of researchers have empirically examined beat perception of musical sequences and related their findings to structural features in the musical stimulus. Drake et al (1997) examined the ability of listeners to tap in time with a variety of piano music by composers such as Bach, Chopin and Bartok. Listeners were required to familiarise themselves with the musical excerpts. They then tapped in synchrony with the music at the most appropriate speed. The majority of listeners responded to mechanical performances at the level of the half measure or measure. Vos et al (1994) examined the ability of trained musicians to perceive the metre of synthesised unaccompanied melodies for cello, violin and flute by Bach. Listeners were instructed to tap to the first beat of every bar in the music. Vos et al (1994) report a greater frequency of tapping responses to positions on the first beat of each bar and a variability in deciding the appropriate length of the bar.

Theoretical work

As discussed above, there have been a number of empirical studies directed towards examining beat and metre perception in basic rhythmical sequences. Very few studies use real music to examine beat perception. There has also been a substantial amount of theoretical work carried out in the area of beat and metre perception. Theoretical approaches have directed attention towards examining how listeners might arrive at a particular metrical interpretation of a given rhythmical sequence. They typically attempt to describe how structural factors, e.g. note duration, in rhythmical sequences may induce a particular metrical interpretation in the listener (Steedman, 1977; Lee 1985; 1991; Longuet-Higgins, 1976; Longuet-Higgins & Lee, 1982). The notes occurring at the beginning of a piece are considered to be most important in determining the perceived metre (Steedman, 1977; Longuet-Higgins & Lee, 1982; Longuet-Higgins, 1976). Steedman's (1977) model of metre perception describes how structural information in the music is used by the listener in order to establish the basic metre. He identifies repetitions of melodic phrases and relatively long notes as the main structural features which determine the listeners metrical interpretation. He argues that the metrical interpretation of piece is founded heavily on early information in the piece and argues that singular challenges to the established metre are non-influential.

Lee (1985) proposes a perceptual model of how the listener infers the metre of simple sequences from information in the relative lengths of notes, i.e. note durations. This theoretical study does not address the role of pitch in the perception of metre. The justification for this serious omission is that information about note length is present in all musical sequences which lend themselves to a rhythmic interpretation. In order to arrive at a detailed account of how listeners perceive the metre of a piece of music it is necessary to understand the effect of note durations in isolation. Once this has been established, pitch factors can then be incorporated into an existing model of rhythm perception. Such an approach is appealing and it seems sensible to begin with this type of reduced stimuli. However, pitch patterns are an essential part of most of the music we listen to. It is most likely that listeners make use of the repeating pitch patterns which many composers include in their music. Lee (1991) extends and describes his model in a more detailed form, incorporating a pre-set tolerance variable which determines whether counter evidence in the sequence will affect the metrical hypothesis.

This results in the generation of more than one interpretation to a given rhythmical sequence. The model also considers the effect of tempo of the sequence in relation to a preferred beat rate which is reported to lie within 0.3 - 0.6 seconds per beat. This kind of model provides a helpful descriptive base on which to relate empirical findings. However, the theoretical study of how duration differences may affect the listeners interpretation of metre, no matter how detailed, will always lack the ability to generalise to real music.

Assessment procedures

An important question in beat perception research is *how should beat-tapping response be assessed?* In other words, how well are listeners tapping to the beat of the music? Empirical studies assess performance by deciding on some criterion for accuracy, for example, listeners are graded as correct if their tapping responses coincide within +/- 10 % of a metrical level such as crotchet or minim beat (Drake et al, 1997). The number of performers tapping at each level is counted to show the proportion response of tapping at each metrical level. Others have graded tapping performance by examining the variability in temporal position of the taps relative to the actual 'blip' in the stimuli (Deutsch, 1983). A number of performance assessments are used in the present study and these are discussed in further detail in the results section below.

Limitations of previous research

Of the empirical studies which examine beat perception, a limited number relate their results to features in a 'real' musical score (Vos et al, 1994) and to accent structures and performance related expressions in the music (Drake et al, 1997). It is clear that this is an area which is in need of further study. In addition to a lack of beat perception research with real music, no experiments to date have examined how long it takes a listener to establish the underlying beat of the music, most preferring to drop data from the first measure (Drake et al, 1997) or to ask participants to wait until they have accurately established the beat before making an accurate response (Vos et al, 1994). This is a serious omission given that establishing the underlying beat is one of the most basic perceptual responses to occur in a music listening experience. In addition to this, theoretical predictions have been made about the time it may take a listener to establish the beat of a sequence. It is suggested that listeners assume that the downbeat occurs on the first note. They will revise this hypothesis if they perceive evidence to suggest otherwise before the first bar has been established. These predictions still need to be empirically tested (Lee, 1991).

What experimental stimuli should be used for beat perception research?

We have already ascertained that a simple finger-tapping response is the ideal task for a beat perception experiment and also that there is a lack of empirical work which uses real music. The present discussion now turns to the question of *what is the most suitable experimental stimulus to use in a beat perception experiment?* In order to answer this question, the aims of the study must first be considered. The aim of the present experiment was to examine the beat perception of music and to relate response patterns to structural features in the music. More specifically, the experiment aimed to:

1. Obtain an objective measure of how long it takes the listener to establish a beat and
2. Examine how much information the listener has obtained from the melody over time by looking at the variability of tapping response (this point is discussed in greater detail in the methods section, below).

Results of these analyses will be related to structural features in the musical stimuli. There is a lack of beat perception research using real music. Given that beat perception is one of the most fundamental responses to most of the music we listen to, this deficit in beat perception research should be addressed. The ideal stimuli should be real music. In addition to this, the music should be unfamiliar and must contain unpredictable rhythm and pitch structures. This is to ensure that listeners actively attend to the music whilst performing the beat-tapping task. Recall that the aim is to relate results to structural features of the music. It is essential the listeners are processing the music as they perform the tapping task. The Gaelic melodies used in the experiments of chapters 2 and 3 are ideal for this purpose. Three Gaelic melodies were used in the present beat tapping experiment. These are presented in detail, below, in the methods section.

In addition to choosing suitable musical material, the music should ideally be presented without expressive timing variations and without expressive dynamic accentuation to minimalise response bias due to a particular performance (Vos et al, 1994). As we have seen in chapters 2 and 3, this is easily done by producing the melodies on computer with accurate timing and without dynamic accents on the first beat of each bar in the music. All notes are played at exactly the same tempo and loudness.

It is not the aim to make the beat perception task artificially difficult so it is important to note that it is easy to tap along to unaccented mechanical piano performances (Drake et al, 1997). Establishing the metre of unaccompanied melodies without timing and expressive deviations is a well-defined and important task of musical understanding (Steedman 1977). It is possible that in using metrically accented stimuli we obtain a measure of the listeners ability to perceive the accented notes in the melody and to synchronise their tapping response with these highlighted metrical positions. Accenting positions in experimental sequences which correspond to the metre in the score could restrict tapping responses to those accented positions alone. As has been noted, the main interest of this study is to allow the listener to respond to their preferred level of the beat.

Experimental hypotheses

It was hypothesised that:

1. The first bar would be the length of time taken to establish an underlying beat. If we consider that evidence from the first bar may be used by listeners to alter their metrical hypotheses (Lee, 1991), then a strong sense of the underlying beat of the melody should be established within the first bar.
2. The majority of listeners would perform with inter-tap intervals which are smaller than the measure level in order to maintain the beat at an accurate rate.

Previous research will clarify hypothesis 2. It has been shown that participants can synchronise their tapping responses with stimuli at rates of between one event per 0.2 - one event per 1.8 seconds with 0.6 seconds as the favoured inter-tap interval for tapping responses (Fraisse, 1982; Parncutt, 1994). At rates faster than one event per 0.2 seconds and slower than one event per 1.8 seconds, tapping responses are far more variable. When listeners synchronise with events which are separated by intervals greater than 1.8 seconds between events, additional memory factors make performance harder (Fraisse, 1982). It has also been demonstrated that beat perception usually occurs in the region of one beat per 0.2 to 1.8 seconds (Parncutt, 1987). Responses at the measure level in the Gaelic melodies used in the present research would require a beat percept every 2.4

seconds for melodies H2 and N1 and 3.6 seconds for melody S1. These inter-tap intervals are far larger than any previously reported. It was therefore predicted that the most likely rate of tapping would be a crotchet or minim beat with inter-tap intervals of 0.6 or 1.2 seconds, respectively. This prediction takes into account the tempo of the present experimental melodies (which is one crotchet beat per 0.6 seconds) and the preferred rate of tapping (Fraisse, 1982; Parncutt, 1994).

For information in the melodies, it was hypothesised that:

1. Maximum information would be communicated as the melody began (Steedman, 1977) and also as a new phrase began in the melody.
2. The first beat of each measure should provide higher metrical information than notes within the bar (Sloboda, 1985; Palmer & Krumhansl, 1990). In other words tapping errors will be more frequent for positions of low metrical information.
3. Uncertainty of beat-tapping response will be minimal at positions in the melody where long, salient notes occur in metrically strong positions (metrically strong positions refer to the first beat of the bar).

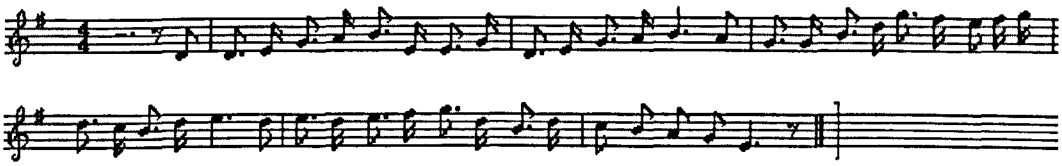
Method

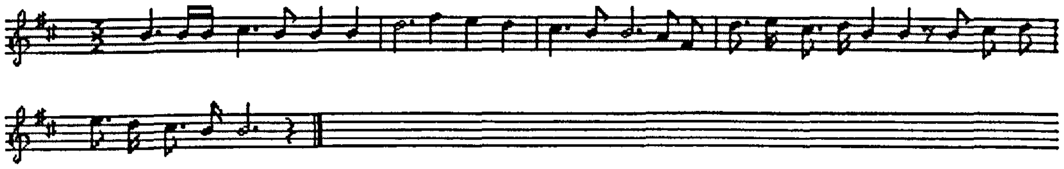
Subjects The subjects were 23 psychology undergraduates, 13 female, 10 male, aged between 17 years and 32 years. Participation for thirteen of the subjects was in part fulfilment of a psychology course requirement whilst the remaining ten were acquaintances of the experimenter.

Stimuli/Apparatus The stimuli consisted of excerpts from three Scottish Gaelic melodies, H2, N1 and S1, shown in Figure 1, below. The melodies were selected from a large set that had been rated on a perceptual-emotional quality scale (Gerstley, 1997). One melody from each of the sad (S1), happy (H2) and neutral (N1) categories was used. Positions of preferred segmentation were determined for each of these melodies in Experiment I, chapter 2. The stimuli served two main purposes:

1. Listeners were unfamiliar with the melodies
2. The rhythmic and pitch structures of these melodies placed suitable demands on the listeners' attention, so a considerable amount of effort was required to follow the underlying beat of the music.

Figure 1. Melodies used in the beat perception experiment.

Melody H2

Melody N1

Melody S1

These were produced by a Yamaha synthesiser with bagpipe timbre on a NeXT computer. Excerpts were played at a tempo of one crotchet beat per 0.6 seconds which is the most natural rate of tapping and the most optimal level for perceptual organisation of a sequence into a series of taps (Fraisse, 1982). The melodies were played without any expressive timing deviations or dynamic variations in order that beat tapping response could be examined in relation to structural features of the music (as discussed above). Listeners responses (onset and duration) were detected by an electronic switch and recorded with a sound edit package on a Macintosh Centris 650.

Procedure A pilot study showed that some listeners experienced difficulty in distinguishing tapping to the underlying beat from tapping the rhythm of the melody. This could result from the listeners unfamiliarity with the melodies. The particularly demanding rhythmic and pitch patterns in the stimuli may distract the listener from maintaining an underlying beat. In addition to this, musically naive participants may find the concept of musical beat difficult to understand. Task explanations were made as clear and informative as possible to ensure listeners understood the task requirements. Listeners were informed that they were required to tap in time to the beat of the music. The distinction between the rhythm of a melody and the beat of a melody was clarified using the analogy of a musician in a marching band. It was pointed out that the musician played a *rhythm* while marching to the *beat.* This technique was used successfully with musically naive participants in a previous beat tapping study (Handel & Oshinsky, 1981). A melody excerpt, similar to the experimental stimuli, was played and the experimenter tapped the rhythm of the melody and the metre of the melody to highlight the difference between these two types of rhythmic task. Once listeners understood the task requirements they were presented with the experimental trials.

Listeners heard the experimental melody once without an active response and then tapped in time with the melody as it was replayed. They tapped their responses onto a hand-held button. The presentation order of the three experimental melodies was counterbalanced across listeners. Listeners completed a short questionnaire which explored musical background and familiarity with the Gaelic melodies. Experimental trials lasted an average of 20 minutes.

Results

The analysis examined several facets of the beat-tapping response.

1. The length of time taken for listeners to establish a regular beat was examined using frequency of correct tapping response across positions in the melody. Responses were considered correct if they were made with an 85 % accuracy in relation to the underlying crotchet beat. The crotchet beat was played at tempo of 0.60 seconds. If a listener was to respond to the crotchet beat of the melody, a correct response would be to tap every 0.60 seconds. It is obvious that not all listeners are going to tap exactly on the beat as they respond, thus a level of 15% error was allowed. This meant that responses which lay within 0.09 seconds before and 0.09 seconds after each actual crotchet beat position were considered correct. Measurements were less stringent for the following analyses and taps were binned into the nearest crotchet beat position for each analysis.
2. Positions at which listeners perceived the beat as occurring in the melody were identified. The accuracy with which listeners could locate the beat was examined using the mean error of tapping response in relation to the nearest crotchet beat of the melody.

3. The length of intervals between tapping responses was examined using inter-tap intervals or distance in time between each consecutive tap. The inter-tap interval determines listeners' consistent tapping rates (Steedman, 1977). Consistency was defined by a particular rate or response which lasted over two measures. Thus, to be marked as a crotchet rate, listeners had to make at least five consecutive crotchet beat responses, to be marked as a minim beat at least three minim beat responses were required, and to be marked as a measure at least two semibreve beat responses had to be made.

4. A measure of information in the melody was calculated. The method for examining the flow of information made use of Fisher's (1950) statistical proposition that the variance of a data set can indicate the amount of information in the data. He proposed the formula $i = 1/nV$ where i = information, n is the number in the sample and V is the variance of the data. This information measure has been used empirically to equate the variability of synchronised tapping responses with information in experimental sequences (Royer, 1967). Others have used it theoretically, to model how sequential stimuli, such as letters, may be processed into progressively larger units of information (Vitz & Todd, 1969).

The basic idea of the present experiment is that listeners' responses will be more certain, i.e. less variable, when beat or metrical information is successfully communicated in the melody. An examination of points of high and low variability in response should show positions of higher and lower beat and metrical information. Patterns of response variability can be considered in relation to structural features in the music.

Responses within the first two phrases of each melody were examined in detail. This length of melody provided an adequate range of positions over which to examine the pattern of information communicated to listeners as the melody progresses. Also, taking responses from the start of the melody onwards enables the calculation of how long it takes to establish the beat of the melody. Responses to the first two phrases of each melody are displayed below in a number of ways. First of all, the time taken to establish the minim beat of the music is shown in Figures 1a, 2a and 3a, below, for melodies H2, N1 and S1, respectively.

As is shown in Figure 1a, it takes 1.5 seconds (3 beats, 6 notes) for 60 % of listeners to establish a minim beat in H2. The highest frequency of response (80 %) occurs when the first long note is sounded (3.9 seconds into the melody).

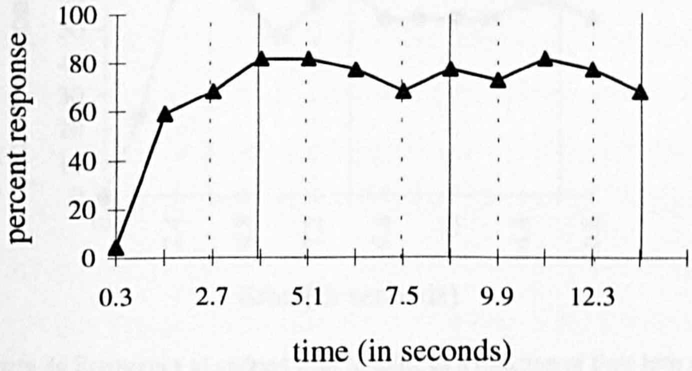


Figure 1a Frequency of correct beat tapping as a function of time into melody H2.
Time into the melody is shown on the x-axis and percent response is displayed on the y-axis. Dotted lines mark the measures in the melody whilst heavy lines mark the onset of the longest notes and phrase boundaries in the melody. The first note of the melody is omitted as responses began on the second note of the melody which is also the start of a measure. A crotchet beat = 0.60 seconds, minim = 1.20 seconds, and a measure = 2.40 seconds in time.

Figure 2a, below, shows a similar pattern of response for N1, with 60 % of listeners establishing a minim beat at 1.8 seconds (4 beats, 5 notes) into the melody. The highest frequency of response occurs just before the occurrence of the longest note (dotted minim) in the melody (6.6 seconds into the melody).

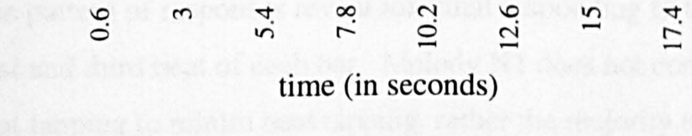


Figure 2a Frequency of correct beat tapping as a function of time into melody N1
Time into the melody is shown on the x-axis and percent response is displayed on the y-axis. Dotted lines mark the measures in the melody whilst heavy lines mark the onset of the longest notes and phrase boundaries in the melody. The first two notes of the melody are omitted as responses began on the third note of the melody.

It takes 2.4 seconds (5 beats, 6 notes) of melody S1 for 60 % of listeners to establish the minim beat (Figure 3a, below). This is as high as the frequency of responses rise for this melody. In the first two phrases of H2 and S1, the highest frequency of response occurs on the long notes in each melody and on long notes occurring on the first beat of a bar.

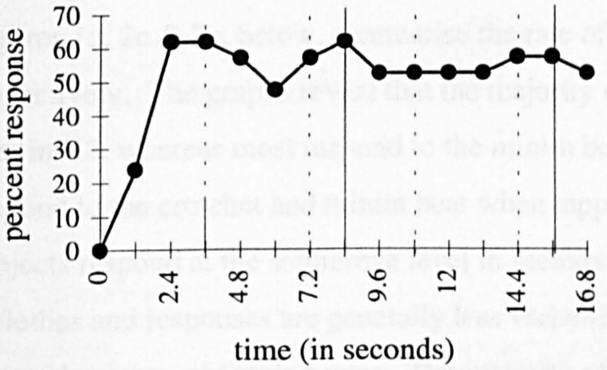


Figure 3a Frequency of correct beat tapping as a function of time into melody S1
Time into the melody is shown on the x-axis and percent response is displayed on the y-axis. Dotted lines mark the measures in the melody whilst heavy lines mark the onset of the longest notes and phrase boundaries in the melody.

Generally, the longest notes elicit the greatest degree of responding in melodies S1 and H2. Responses to N1 are more consistent across different melody positions with only one high frequency of response (80 % response 6.6 seconds into the melody). This does not occur on an important metrical position or on a relatively long note.

The less stringent analysis (as discussed above) which bins all beat tapping responses into the nearest crotchet beat is used for the remaining analyses. The frequencies of beat tapping response to the nearest crotchet are shown in Figures 1b, 2b, & 3b, below, for melodies H2, N1 and S1, respectively. The general response patterns of H2 and N1 (Figures 1b and 2b, respectively) fit in with the underlying metre as indicated on the musical score. There is an initial high response at the crotchet beat level in H2 in the first phrase followed by a greater degree of minim beat responding in the second phrase. The pattern of responses reveal maximal responding to those beats which occur on the first and third beat of each bar. Melody N1 does not contain this switch from crotchet beat tapping to minim beat tapping, rather the majority of listeners establish the minim beat rate in the first measure of phrase one. The pattern of response is similar for the repeated phrase in this melody (the repeated section is marked by the bold line in the centre of Figure 2b). Again, listeners tend to tap more often to those beats on the first and third beats of the bar. There is a similar response pattern for melody S1 in that listeners initially respond to every crotchet beat of the bar then reduce their tapping rate to a minim beat. The highest frequency of response coincides with the first, third and fifth beat of each bar.

Figures 1c, 2c & 3c, below, summarise the rate of response tapping for H2, N1 and S1, respectively. The graphs reveal that the majority of listeners respond to the crotchet beat in H2, whereas most respond to the minim beat in S1 and roughly equal numbers respond to the crotchet and minim beat when tapping to N1. A greater percentage of subjects respond at the semibreve level in melody H2 in comparison with the other two melodies and responses are generally less variable for H2 than N1 and S1 which result in a wider range of tapping rates. The majority of listeners were consistent in their tapping rate after the five few seconds of hearing the melody. Only a small number of listeners did not manage to establish a regular beat at any level and altered their tapping rate continuously as the melody progressed.

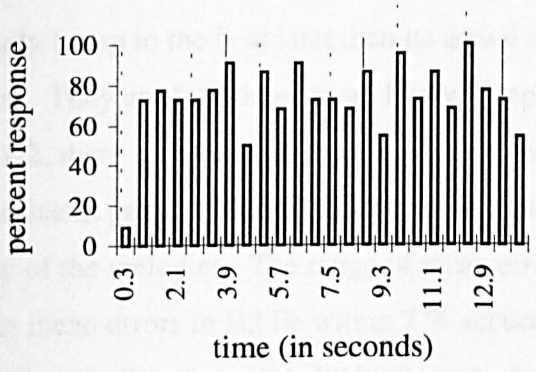


Figure 1b Frequency of tapping response in H2
Time of the melody is shown on the x-axis and percent response is displayed on the y-axis. Dotted lines mark the measures in the melody with heavy lines marking the onset of the longest notes in the melody.

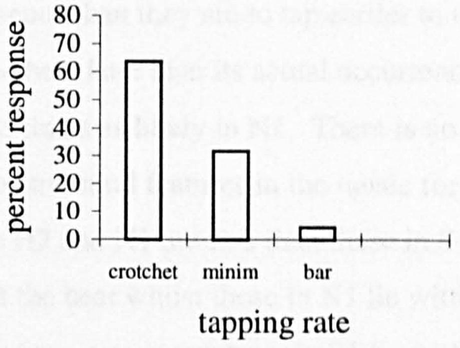


Figure 1c Frequency of consistent tapping rate in H2
Percentage of listeners responding at a particular rate is shown on the y-axis. A crotchet beat = 0.60 secs, a minim = 1.20 secs, and a measure = 2.40 secs in time.

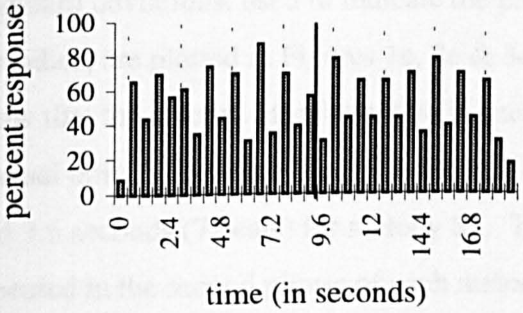


Figure 2b Frequency of tapping response in N1
Time of the melody is shown on the x-axis and percent response is displayed on the y-axis. Dotted lines mark the measures in the melody with the heavy line marking the start of the repeated section in the melody.

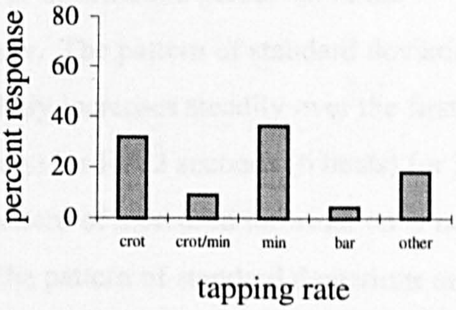


Figure 2c Frequency of consistent tapping rate in N1
Percentage of listeners responding at a particular rate is shown on the y-axis. A crotchet beat = 0.60 secs, a minim = 1.20 secs, and a measure = 2.40 secs in time.

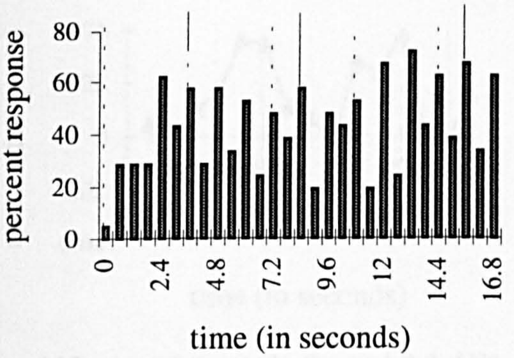


Figure 3b Frequency of tapping response in S1
Time of the melody is shown on the x-axis and percent response is displayed on the y-axis. Dotted lines mark the measures in the melody with the heavy lines marking the phrase boundaries in the melody.

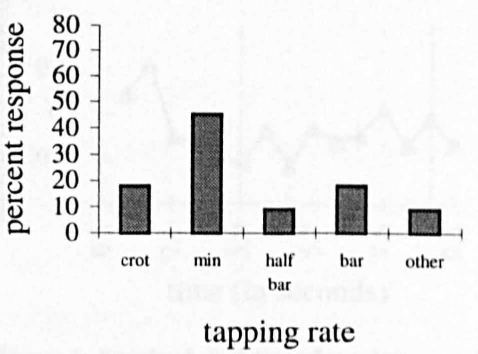


Figure 3c Frequency of consistent tapping rate in S1
Percentage of listeners responding at a particular rate is shown on the y-axis. A crotchet beat = 0.60 secs, a minim = 1.20 secs, and a semibreve = 2.80 secs in time.

The mean errors of the first two phrases are shown in Figures 1d, 2d & 3d, below. Overall, there are more late errors than early errors for the melodies. Listeners are more likely to tap to the beat later than its actual occurrence than they are to tap earlier to the beat. They are four times more likely to tap to the beat later than its actual occurrence in H2, three times as likely in S1 and one and half times as likely in N1. There is no consistent pattern of error which can be related to structural features in the music for any of the melodies. The range of mean errors in H2 and N1 are less than those in S1. The mean errors in H2 lie within 7 % accuracy of the beat whilst those in N1 lie within 8 % (excluding the unusually high error of the first response) and those in S1 lie within 11 % of the beat. There is a trend for more accurate responses as melody N1 progresses. This trend is not paralleled in H2 and S1.

Standard deviations, used to indicate the pattern of information perceived in the melodies, are plotted in Figures 1e, 2e & 3e, below. The pattern of standard deviations show that the amount of information in each melody increases steadily over the first phrasal unit. This occurs over 3.9 seconds (7 beats) for H2, 3 seconds (6 beats) for N1 and 3.6 seconds (7 beats) for melody S1. This pattern of increased information is not repeated in the second phrase of each melody. The pattern of standard deviations are more variable as the melody progresses into phrase 2. It is noteworthy that the largest standard deviations in the second phrase of N1 and S1 occur on the note directly after the long held notes in the melodies.

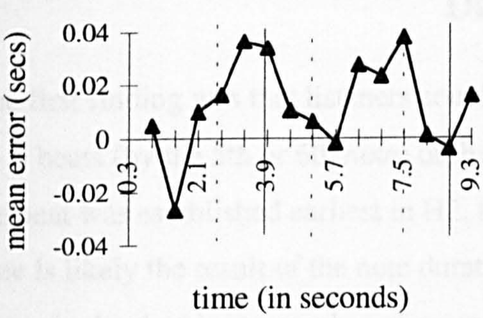


Figure 1d Beat tapping error in phrases 1 & 2 of H2
Mean error is displayed on the y-axis. A negative score is an early error (tapping responses are earlier than the beat), a positive score is a late error and zero is correct (listeners respond on the actual beat).

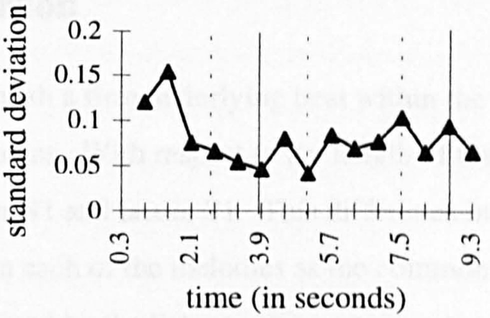


Figure 1e Standard deviation of tapping error in H2
Standard deviation of beat tapping response is shown on the y-axis. Data for the first beat are omitted due to the small number of listeners responding at this point.

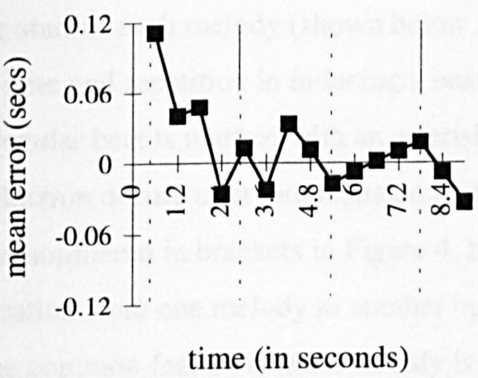


Figure 2d Beat tapping error in phrase 1 - 2 of N1
Mean error is displayed on the y-axis. A negative score is an early error (responses are earlier than the beat), a positive score is a late error and zero is correct (listeners respond on the actual beat).

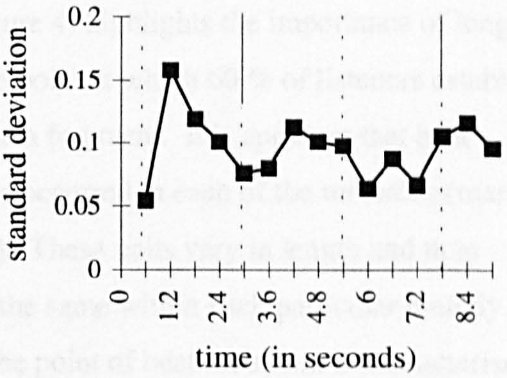


Figure 2e Standard deviation of tapping error in N1
Standard deviation of beat tapping response is shown on the y-axis. Data for the first beat are omitted due to the small number of listeners responding at this point.

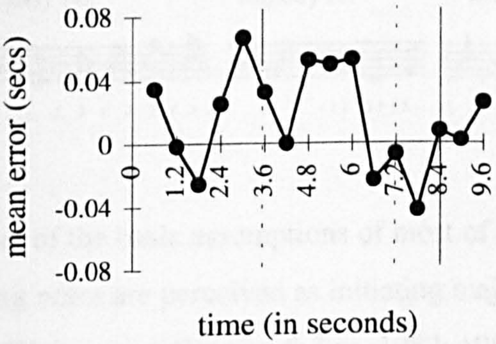


Figure 3d Beat tapping error in phrase 1 - 2 of S1
Mean error is displayed on the y-axis. A negative score is an early error, (responses are earlier than the beat), a positive score is a late error and zero is correct (listeners respond on the actual beat).

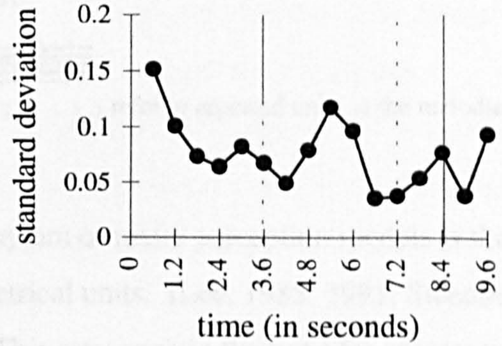
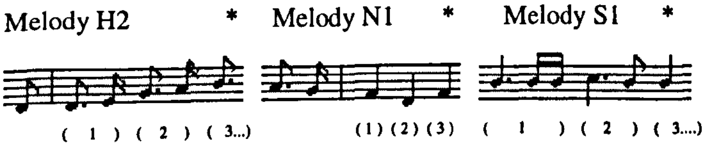


Figure 3e Standard deviation of tapping error in S1
Standard deviation of beat tapping response is shown on the y-axis. Data for the first beat are omitted due to the small number of listeners responding at this point.

Discussion

The first finding was that listeners could establish a firm underlying beat within the first 3 - 5 beats (by the 5th or 6th note) of the melodies. With respect to the length of time, the beat was established earliest in H2, then in N1 and last in S1. This difference in time is likely the result of the note durations in each of the melodies as the common factor in the data is the *number of notes* perceived by the listener. Five or six notes were required to be heard in each melody before the beat was established. This is regardless of the overall time into each melody. It seems, therefore, that the number of sounded events is one of the main factors which effects the perception of metre or induces a stable beat in response to these melodies. A closer look at the fragment of the start of each melody (shown below in Figure 4) highlights the importance of length of note and repetition in inducing a beat. The point at which 60 % of listeners establish a regular beat is marked with an asterisk in each fragment. It is apparent that beat induction occurs after two repeated units have occurred in each of the melodies (marked and numbered in brackets in Figure 4, below). These units vary in length and note duration from one melody to another but are the same within each particular melody. The common feature of each melody is that the point of beat induction is characterised by the repetition of three groups of salient notes. The dotted quavers in H2, crotchets in N1 and dotted crotchets in S1 are the salient notes.

Figure 4. Length of melody required to establish a beat tapping response.

refer to repeated units in the melodies.

One of the basic assumptions of most of the rhythm or metre perception models is that long notes are perceived as initiating major metrical units. (Lee, 1985; 1991; Steedman, 1977; Longuet-Higgins & Lee, 1982; 1984). This may explain the need for at least two relatively long notes in communicating the underlying beat of the music. Lee (1985) suggests that repetition is an important factor even in the interpretation of impoverished sequences. Steedman (1977) considers that repetition may cause the listener to alter his metrical interpretations. This may be because listeners are cautious and are not therefore likely to alter the metrical framework they have been using to successfully track the music. An isolated case which calls the established metre into question is unlikely to initiate a change in metrical interpretation. However, repetitions which continuously question the established metre will lead to the realisation that the music is not being successfully tracked and listeners may eventually alter their metrical interpretation of the piece. The present results support the idea that listeners use note lengths or salient notes and repetitions of these notes when establishing a beat. Two groups of repeated longer notes in the melody successfully induce an underlying beat. The results also tie in well with evidence which suggests that synchronisation is established fairly early in a sequence of sounds. Tapping responses are generally synchronous by the third pattern in sequences of repetitive patterns (Fraisse, 1982). This finding is reflected in the above results.

Long notes and accented notes enhance metrical structure and are perceived as initiating higher level metrical units whereas shorter or unaccented notes are not (Lee, 1985). This may go some way to explaining the rate of tapping response in the melodies. The preferred rate of tapping seems to be dependent on structural features of the melody. Most listeners chose to synchronise with the crotchet beat in H2, equal numbers chose the crotchet and minim beat in N1 and most chose to respond to the minim beat in S1. Lee (1985) argues that listeners use repetitions of melodic fragments as determinants of the metrical unit by taking the distance between the starting point of one unit and the starting point of a repeated unit to give the metrical unit. The difference between the predominant crotchet beat responding in H2 and minim beat responding in S1 can be explained by considering this idea of repeating melodic fragments as initiating a metrical unit. A common repeating pattern in H2 was the dotted quaver and semiquaver fragment lasting a crotchet beat. Most listeners used the crotchet beat as their response rate in H2. The most common repeating fragment in S1 was the dotted crochet and quaver fragment which corresponds to a minim beat. The rate at which most listeners responded in S1 was the minim beat. The repeating fragment in N1 was the single crotchet and equal numbers of listeners responded to minims and crotchets.

As was expected, most listeners generally respond at either the crotchet or minim beat level, in line with previous studies which examined preferred rates of tapping response

(Fraisse, 1982; Parncutt, 1994). Very few listeners responded at the level of the measure in the present experiment. The frequency of beat-tapping responses suggest that many listeners perceive the melodies to be in 2/4 metre, with 32 % of listeners responding to the minim beat in H2, 36 % responding to the minim beat in N1 and 45 % responding to the minim beat in S1. Given the time signature of melodies H2 and N1, responses to the minim beat are responses at the half measure level. This is in line with results of musicians reported by Vos et al (1994). They found that musicians tend to tap to the beat at the half measure aswell as at the measure level. Very few listeners pick up on the 3/2 metre in S1, most choosing to respond at the minim beat level. Minim beat responses in S1 enable the listener to tap at important metrical positions such as the first beat of each bar.

A wider range of tapping was observed in N1 and S1, suggesting a larger degree of uncertainty as to which was the metrical level (Drake et al, 1997). If we examine structural features of the melodies, we see that the main difference is the greater length of long held notes in N1 and S1 compared with H2. It may be that metrical information decreases or is more ambiguous after these long notes have occurred as there are no sounded events to mark the beat. The success of listeners to respond to beats at the start of a measure and to long notes which coincide with the first beat of a measure may be due to the salience of the notes at these positions. This salience may in turn attract attention towards the metrical information at these positions. However, metrical information seems to decrease immediately after a long note in the melodies during which time there are no crotchet beat markers (e.g. a dotted minim).

The mean errors highlighted a response bias towards late rather than early responding to the beat. This is in contrast to previous research on motor responses to periodic stimuli which reports a general trend for early synchronisation in the region of 0.03 seconds (Boring, 1950; Fraisse, 1987). Early responses are explained within an anticipation framework (Boring 1950) and by the time delay between movement of the hand and the central perception of this movement in the brain (Fraisse, 1987). It is argued that reactions which precede the stimulus highlight the fact that the signal for the response is not the actual sound of the melody itself but the temporal interval between each of the beats in the music. The very nature of synchronising a response to music necessarily contains a strong element of anticipation and can occur when the sequence is regularly structured in a clear and well defined manner. The consistent late responding found in the present experiment suggests that, rather than mindlessly anticipating the beat, listeners were attending to the musical structure as the melody progressed. We would expect a latency of response in relation to the on-going beat in cases where listeners attend to the music and tap with the beat rather than anticipate the beat. This account seems likely if we consider that the sequences which attracted the early responses in previous studies consisted of non-musical click sounds. Once the beat is established in these sequences there are no varying durations or unexpected pitch patterns to throw the listener. Sequences of equally spaced clicks are completely predictable and as a result are more likely to be easier to track than, for example, an unfamiliar Gaelic melody. There is a considerable degree of effort required to synchronise a tapping response with melodies which contain unfamiliar pitch and rhythmic patterns. Listeners have to attend to the melody primarily which may account for the tendency for late rather than anticipatory beat tapping responses.

The measure of information in the melodies revealed that maximal information was communicated within the first phrase of each melody as standard deviations decreased steadily throughout the first phrase. In phrase two, standard deviations were variable with no consistent pattern of information flow in relation to the musical structure. The higher variation in tapping response to notes which were directly after a long held note in melodies N1 and S1 suggests that metrical information decreases after a long held note has occurred. It is established that beats are perceived when some elements of the pattern are missing, thus responses correspond to actual events and missing events such as perceived beats on the longer notes of the melodies (Parncutt, 1987). What has not been documented is the effect of elapsed time on the ambiguity of beat position. Without an external sounded event with which to synchronise, beat tapping responses may become increasingly inaccurate. The present results suggest that, after a long note, metrical information in a melody decreases. The two melodies which contain longer held notes (N1 and S1) are metrically more ambiguous, shown by a wider range of tapping rates, than melody H2 which does not contain such long notes.

The present results were obtained with perfectly timed and metrically unaccented melodies so any perceptual effects reported can be attributed to the interpretation of information in the score. We can conclude that the pitch and rhythmic structures of these three melodies are sufficient to communicate metrical information to the listener in as little as two repeating patterns of each melody. Moreover, the first phrase of each melody communicates the highest amount of information. These results go some way to explaining why information at the beginning of a melody counts for more than information later on in a melody.

Chapter 5 Perception of musical intervals and Gaelic melodies

It has been demonstrated that it is possible to identify some kind of unit of information in Gaelic melodies to which listeners are sensitive (Chapter 2). In addition to this, an examination of memory for the rhythmic structure of these melodies (Chapter 3) reveals a memorable unit of information which is shorter than the phrasal or segmental unit identified in the Segmentation Experiment. Furthermore, beat induction of these melodies can occur after hearing as little as 5 or 6 notes of the melody (Chapter 4). The next question which this research leads on to is *what is the smallest musical unit to generate a possible meaning?*

Background

That age old question which has eluded musicologists and psychologists alike *what is the meaning of music?* will, unfortunately, not be answered in this final chapter. The main focus of the first section of this chapter is to introduce the final experimental task, the two alternative-forced choice task (2AFC). Previous research using experimental procedures to describe the meaning of music will be considered and a method developed by Watt (1995) will be described in detail.

Various techniques have been developed to examine emotional responses to and perceived meanings of music. The free response method typically requires listeners to write down the adjectives which best describe the music they hear. A number of researchers have attempted to work within a more structured framework, for example, presenting listeners with a list of adjectives which are arranged into groups according to the feelings of the adjectives. Listeners then choose adjectives which are most appropriate to the music they hear (Hevner, 1936; Gabrielsson, 1973). Some researchers argue against the use of such techniques, claiming that although music has characteristics which can be easily identified by listeners, these are poorly described by the adjectives we use (Pratt, 1968). In addition to this, others, such as Meyer (1956), believe that free verbal responses are strongly personal to the participant and are also affected by a multitude of non-emotional factors.

It is widely accepted that the lack of a standardised method for measuring meaning, e.g. of words or music, is problematic. Given the extremely difficult task of developing such a method, there have been few attempts to do so. However, early attempts were made to examine the meaning of words which led to the development of a semantic differential task (Osgood, Suci & Tannenbaum,1957). The task consists of a number of bipolar adjective scales with 7 point rating forms, for example, **Small [ ] [ ] [ ] [ ] [ ] [ ] [ ] Large**. Participants are presented with a stimulus, e.g. a word, and respond by ticking one of the boxes from left to right box which they feel best represents the given word. They perform this task using all the adjective scales which are present. Another word is then presented and so on. One of the main ideas behind this task is to use a factor analysis to determine the main dimensions along which meaningful reactions to different stimuli vary. Those who have made use of the semantic differential scale to describe the meaning of words include Osgood (1952); Osgood & Suci (1955) and Osgood et al, (1957). Others have used the semantic differential task to examine the emotional expression of music (Gfeller, Asmus & Eckert, 1991; Wedin, 1972) and the experience of different rhythmic patterns (Gabrielsson, 1973).

Watt (1995) developed a 2AFC task in an attempt to examine the listeners perception of meaning in unfamiliar musical excerpts. Watt's 2AFC task differs from previous semantic differential techniques by providing no neutral category. The task list contains fourteen adjective pairs, roughly opposites, and forces listeners to choose one adjective or another from *every* pair used in the list. Listeners cannot indicate the degree of their response, unlike semantic differential scales. The rationale behind this forced response is that some of the adjective pairs used in the experiment, such as *male/female*, may not seem like natural categories to select in response to musical stimuli, whereas adjective pairs, such as *joyful/sad*, are more commonly associated with music. If the option of a *don't know* or *neutral* response is offered then listeners may always choose this category in response to categories which are not normally associated with music. If forced to choose, however, it is quite feasible that a significant proportion of listeners will respond by selecting the same adjective from an unusual pair. Indeed Osgood & Suci (1955) report that the extremes of certain items in their 7-point scales are used more often than others and also that ambiguous or unfamiliar concepts tend to be allocated positions near the neutral point in the scale. Watt (1995) eliminates this potential response bias by forcing listeners to choose one or other of the adjective pairs for every type of category used in the experiment. In addition to this, the forced choice response was deemed to be less susceptible to situational biases, such as desire to respond in the way which listeners perceive will be pleasing to the experimenter and prior knowledge of what the music is most likely to mean, which may occur in experiments using a less rigid response paradigm.

The basic assumption underlying Watt's (1995) approach is that music has an ability to express and this serves a communicative function if the expressions are interpreted similarly by most listeners. It is argued that if music is successful in its communicative purpose, then the majority of listeners should be in agreement about descriptions of that music. The agreement or consistency of listeners responses can be calculated (this is discussed in detail in the Results section, below) and this gives a direct measurement of the information content in each musical excerpt. The adjective categories which result in the highest levels of agreement in response are those which recover the highest information content from the musical excerpts.

The types of description on which listeners agreed were examined and this was done for a variety of different types of music ranging from Purcell to (the artist formerly known as) Prince (Watt & Ash, 1996). Having established which types of description are consistently used, these can be considered as psychologically appropriate descriptions of the music. The main findings are that adjective pairs which elicit the highest levels of agreement in response to music are from person categories. In other words, they are those adjectives from the list which can be used to describe people such as female/male and pleased/angry. Non-person categories refer to those adjectives which are not typically used to describe person like attributes such as day/night and bright/dull. The information content recovered from person categories is significantly higher than that of non-person categories for music. The person - non-person effect was not found in response to non-musical control stimuli (four different wholemeal foods). The information content recovered by person categories in music is significantly higher than that of control stimuli. There are no differences between information content recovered from non-person categories in response to music and control stimuli. Watt (1995) theorises that music generates representations which contain similar attributes to those which we assign to people in our everyday human interactions.

Watt's (1995) novel technique offers a non-direct way to examine meaning in musical excerpts which is ideally suited for non-musical participants. As mentioned in chapter 1, in order to examine perceptual responses to music, listeners must make a response which is related to the music in some way. The present 2AFC task obtains a set of such responses in as non-biased a manner as is possible. The task instructions do not direct task performance in any particular way. Listeners are asked to listen to the music then perform the task. They are unaware that the task has anything to do with content and meaning of music.

The results of these 2AFC experiments were obtained using a wide variety of musical excerpts with a multitude of varying factors such as timbre, tempo and style (Watt, 1995; Watt & Ash, 1996). There is no evidence, as yet, to suggest that the same technique would be of use in examining meaning in stimuli which are more easy to dissect, e.g. musical intervals and unaccompanied melodies. Before using the 2AFC task to examine the information prevalent in the unaccompanied Gaelic melodies, responses to melodic two-note intervals of the octave were examined. Intervals are useful musical material with which to examine perceptual responses and have been studied previously both theoretically and empirically in a number of ways (Shepard, 1982; Burns & Ward, 1982; Dowling & Harwood, 1986; Krumhansl, 1990). Given that two notes are the smallest number possible to create a musical interval in a melody, musical intervals are a good test stimuli for the present novel 2AFC task. If results found in response to musical intervals are similar to those found in response to large scale musical forms, then it is likely that the technique will be of use in examining responses to short unaccompanied melodies.

Experiment I Perception of Musical intervals

Experiment I examines what the 2AFC task can tell us about perceptual responses to melodic musical intervals in isolation, i.e. two note melodic intervals are presented on their own and are not presented as part of a larger musical context. The results are considered in relation to the results obtained by Watt (1995) to examine the efficacy of using the 2AFC task with unaccompanied melodies. In addition to this, repeating rhythm patterns are also examined using the 2AFC task to examine the effect of increasing the length of simple musical stimuli on perceived information content. Responses to the rhythm patterns are compared with responses to the two note intervals.

Method

Subjects The subjects were 70 psychology undergraduates, 40 male and 30 female, aged between 17 and 36 years. Participation in the experiment was in part fulfilment of a psychology course requirement. All subjects were non-musicians.

Stimuli The twelve intervals of the octave were used:

minor 2nd
major 2nd
minor 3rd
major 3rd
perfect 4th
augmented 4th
perfect 5th
minor 6th
major 6th
minor 7th
major 7th
octave

Each interval occurred in four different ways:

1. Rising from middle C e.g. C_4 - G_4
2. Falling to middle C e.g. G_4 - C_4
3. Rising to upper C e.g. G_4 - C_5
4. Falling from upper C e.g. C_5 - G_4

There are thus a total of 48 different stimuli. The tempo of each interval was one note (crotchet) per 0.60 seconds.

The following eight repeating rhythm patterns were used:

Note Crotchet = 4, dotted quaver = 3, quaver = 2, semi-quaver = 1

1. 44444
2. 2222222222
3. 211211211211211
4. 112112112112112
5. 121121121121121
6. 1313131313
7. 11111111111111111111
8. 3131313131

These were played on either a C Major chord (C4, E4, G4 & C5) or a single note, C4 with piano timbre using *Deluxe Music Construction Set* on a Macintosh Centris 650. There was at least twelve harmonics in each tone.

Procedure

The following adjective pairs were used in a two alternative forced choice task:

Figure A Sample adjective list used in the 2AFC task

female	☐	male	☐
joyful	☐	sad	☐
old	☐	young	☐
bright	☐	dull	☐
prickly	☐	smooth	☐
narrow	☐	wide	☐
angry	☐	pleased	☐
good	☐	evil	☐
leaden	☐	weightless	☐
sweet	☐	sour	☐
dry	☐	moist	☐
stable	☐	unstable	☐
day	☐	night	☐
gentle	☐	violent	☐

Listeners were required to select one adjective from each pair of adjectives (listeners therefore made fourteen 2AFC responses) after they heard each stimulus. The response categories were chosen to offer a broad spread of possible judgements. The order of the adjective pairs on the page was fully randomised. Intervals were heard once on a standard Sony Walkman. The musical intervals were randomly ordered with verbal instruction in between every interval. Forty-two of the listeners performed with the intervals rising to and falling from C5 whilst the remaining twenty-eight performed with the intervals rising from and falling to C4. Listeners controlled the rate at which they performed the experiment. The experiment lasted an average of 40 minutes.

Analysis

Two main analyses were carried out on the data.

1. A factor analysis was performed on the mean response to each category for each interval to establish on how many dimensions the perception of musical intervals vary. Mean response was calculated by counting the number of selections of a given adjective from each pair in response to each stimulus. The count was divided by the number of participants to give a proportion response.

2. The proportion of listeners responding with the adjective on the right of each pair in the above example (Figure A) were counted separately for every musical interval. A

measurement of *information content* was calculated for each musical interval and each adjective pair using the following formula:

information content = abs(proportion response - 0.5) X 2

This results in a linear measure of the deviation of response from chance which ranges on a scale from 0 - 1. The information measure takes into account consistency of response whilst ignoring the direction of response. A few examples will clarify the use of this information scale which basically indicates consistency of listener response in the 2AFC task. A performance at chance in this task would result in half of the listeners selecting the box on the left and the other half selecting the box on the right. The proportion response to this adjective category would be 0.50. Using the formula above,

information content = abs(0.50 - 0.50) X 2

the information content in this particular example is *0*. A more consistent response in one direction would result if a greater proportion of the listeners chose one adjective over another e.g. a proportion response of 0.75. Using the above formula this would result in an information content of 0.50. A completely consistent response is one in which all listeners select the same adjective from a pair. This results in a proportion response of 1.00. The resulting information content is 1.00 which is the highest value possible on the information content scale.

Results

Factor analysis

The factor analysis was run on a 14 (number of adjective pairs) by 48 (number of interval stimuli) matrix of proportion response. This revealed that there are three main orthogonal dimensions to the perception of musical intervals. In other words, listeners' responses to the two note intervals can be described in three main independent ways. A plot of the eigenvalues (Figure 1) demonstrates that there are three main factors which account for most of the variance of response. Using the Kaiser (1960) test, which works on the basis of dropping those factors which account for less than the equivalent of one variable, three eigenvalues (factors with variances) greater than 1 can be identified.

Factors 1 and 2 account for over 77 % of the variability and will be discussed in greater detail, below. Factor 1 and 2 scores for each of the four interval conditions (ascending from C4, descending to C4, ascending to C5 and descending from C5) are shown in Figures 1a - 2d, below.

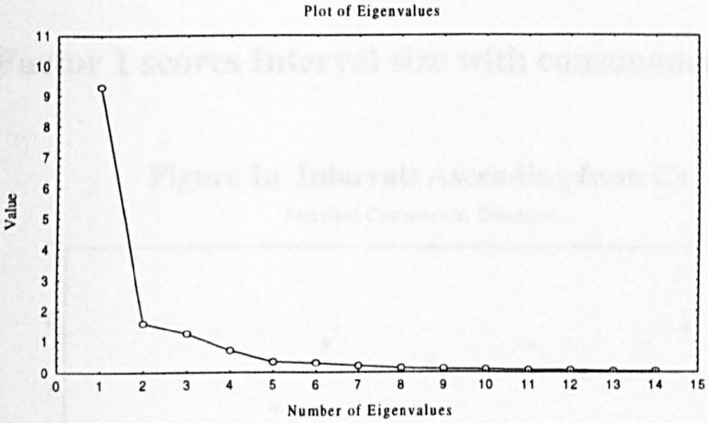


Figure 1 Scree plot of principal components analysis for interval data.

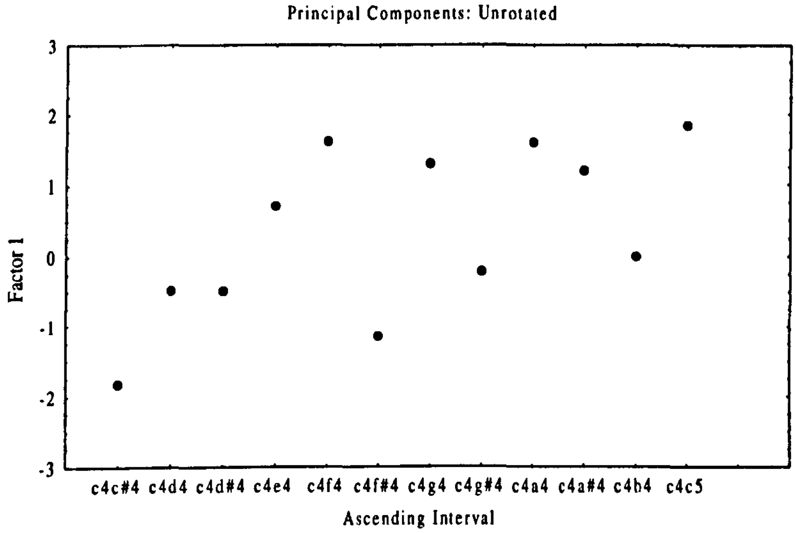
Factor 1

The first factor accounts for 66.28 % of the variability in response. This factor can be interpreted as representing a combination of the listeners' sensitivity to pitch interval size (the intervals ranged from minor second up to an octave) and to the harmonic relationships between certain intervals in the octave. There is a general trend of lower factor scores for the smallest interval for intervals rising from and falling to C4 (Figures 1a & 1b, below). The factor scores increase with an increase in distance between the first and second tone. There is a similar trend, but in the opposite direction, for intervals rising to and falling from C5. Factors scores are highest for the smallest interval and decrease gradually with increasing interval size (Figures 1c & 1d, below).

Figure 1a shows a strong effect of pitch interval size with increasing factor scores as the distance between the first and second note in the interval increases, e.g. minor 2nd is distinct from minor 3rd which is distinct from perfect 4th and so on. In addition to this, positive factor scores occur for the following intervals: major 3rd, perfect 4th, perfect 5th, major 6th, minor 6th and octave. These intervals are considered to be consonant in modern music theory (Sadie, 1981). Figure 1b shows a similar, more gradual, effect of pitch interval size. The highest factor scores occur for major 3rd, perfect 4th, perfect 5th, major 6th and octave. Figures 1c and 1d also show a gradual effect of pitch interval size, with a decrease in factor scores as the interval between the two tones increases. The factor scores corresponding to the perfect 4th, perfect 5th and octave do not decrease with the rest of the intervals in Figure 1c and the factor scores of the perfect 4th and perfect 5th remain high in Figure 1d.

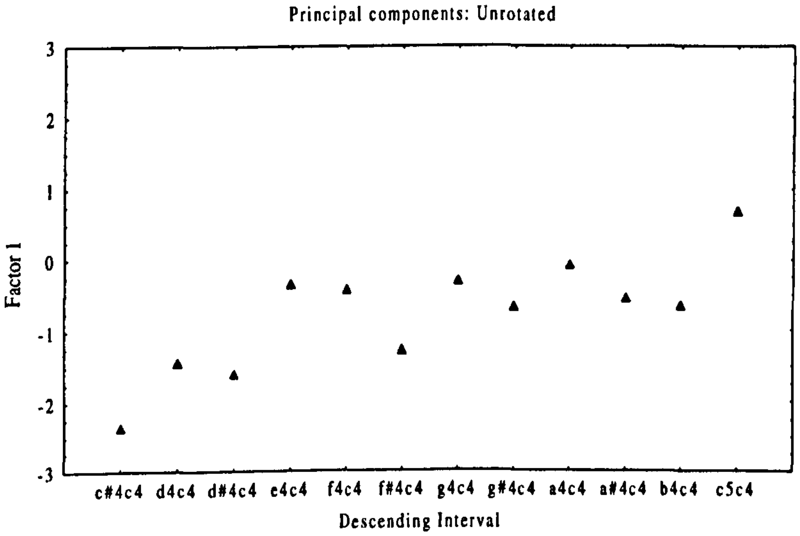
Factor 1 scores Interval size with consonance

Figure 1a Intervals Ascending from C_4



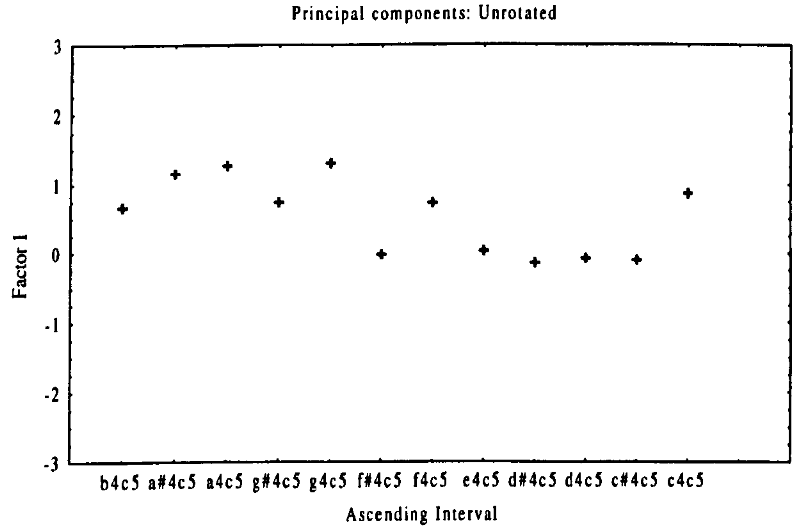
The y-axis shows the factor scores whilst the x-axis displays the intervals which range from minor 2nd to octave. The first and second tones of each interval are described on the x-axis. For example, c4c#4 refers to the interval which contains c4 as the first tone and c#4 as the second tone.

Figure 1b Intervals Descending to C_4



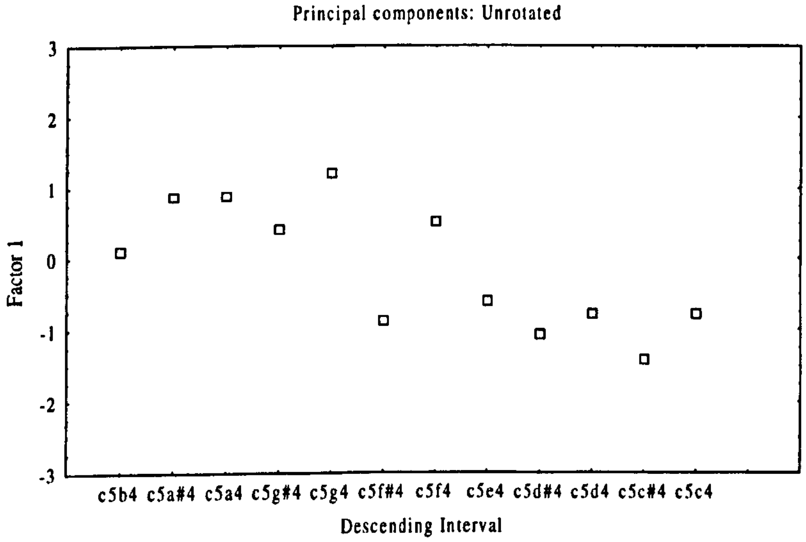
The y-axis shows the factor scores whilst the x-axis displays the intervals which range from minor 2nd to octave. In the case of c#4c4, the first tone of the interval was c#4 and the second tone was c4.

Figure 1c Intervals Ascending to C5



The y-axis shows the factor scores whilst the x-axis displays the intervals which range from minor 2nd to octave. In the case of b4c5, the first tone of the interval was b4 and the second tone was c5.

Figure 1d Intervals descending from C5



The y-axis shows the factor scores whilst the x-axis displays the intervals which range from minor 2nd to octave. In the case of c5b4, the first tone of the interval was c5 and the second tone was b4.

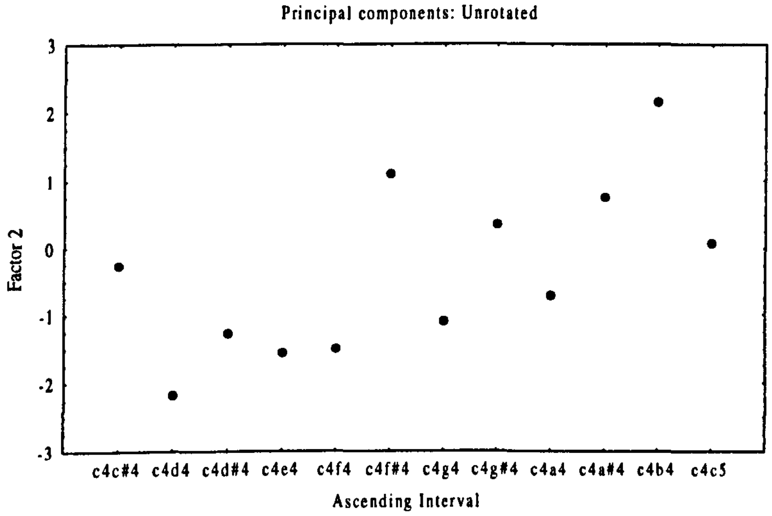
Factor 2

The second factor, which accounts for 11.20 % of the variability, is described generally as *small and large intervals*. This factor is interpreted as representing the listeners binary perception of small conjunct intervals as distinct from larger disjunct intervals. In Figures 2a and 2b, lower factor scores result from the smaller more melodic intervals such as major and minor 3rd & perfect 4th. The higher factor scores occur in response to the larger, less melodic intervals e.g. augmented 4th and major 6th. The change in factor

2 scores is not gradual as is observed in Figures 1a - 1d. The clearest examples are shown in Figure 2b. Scores can be split into three distinct groups dependant on the size of the musical interval. The first group includes those intervals up to perfect 4th, the second group includes those from an augmented 4th up to a major 6th and the third group includes greater than a major 6th. Figures 2c and 2d do not reflect this pattern as clearly.

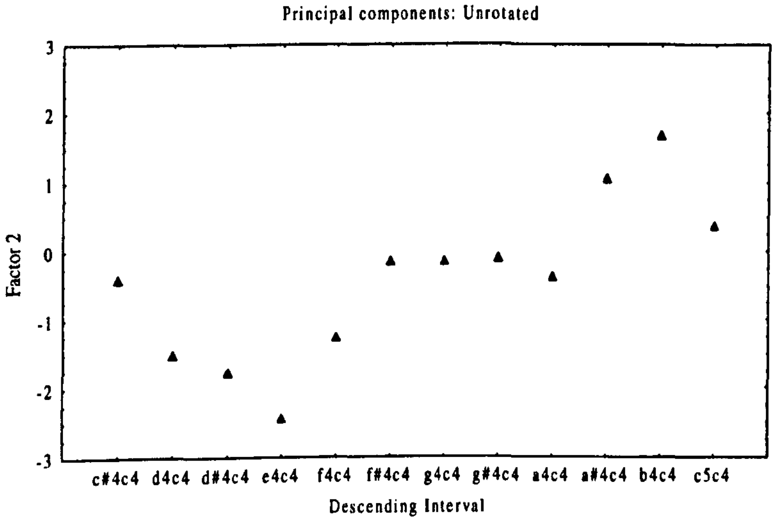
Factor 2 scores Small and Large intervals

Figure 2a Intervals Ascending from C$_4$



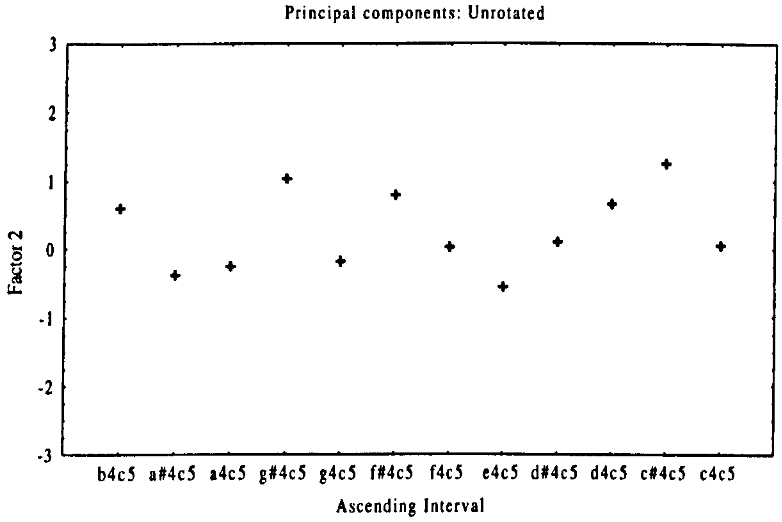
The y-axis shows the factor scores whilst the x-axis displays the intervals which range from minor 2nd to octave. The intervals are described as above in Figures 1a.

Figure 2b Intervals Descending to C$_4$



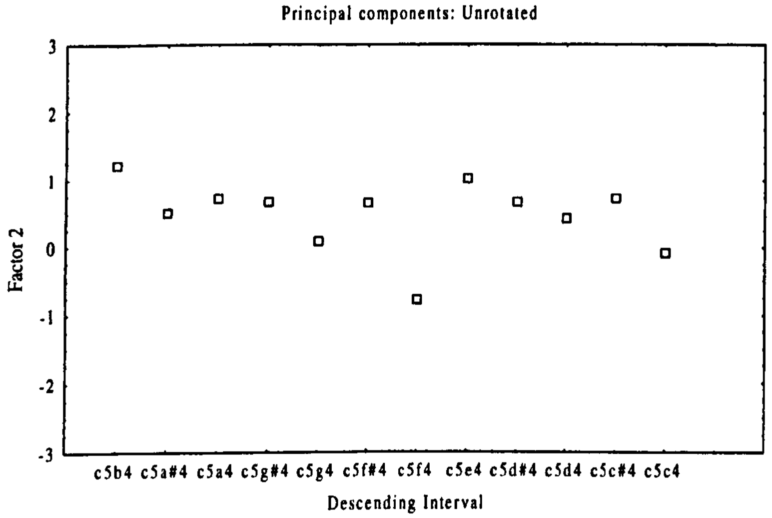
The y-axis shows the factor score whilst the x-axis displays the intervals which range from minor 2nd to octave. The intervals are described as above in Figures 1b.

Figure 2c Intervals Ascending to C_5



The y-axis shows the factor score whilst the x-axis displays the intervals which range from minor 2nd to octave. The intervals are described as above in Figures 1c.

Figure 2d Intervals Descending from C_5



The y-axis shows the factor score whilst the x-axis displays the intervals which range from minor 2nd to octave. The intervals are described as above in Figures 1d.

Information analysis

A comparison of person with non-person categories (which were found to yield different information content values in response to the musical stimuli examined by Watt (1995)) was performed on the interval data.

non-person	*person*
narrow/wide	pleased/angry
dry/moist	evil/good
bright/dull	gentle/violent
day/night	female/male

The results of this comparison for the four interval groups used in the experiment are displayed in Figure 3, below.

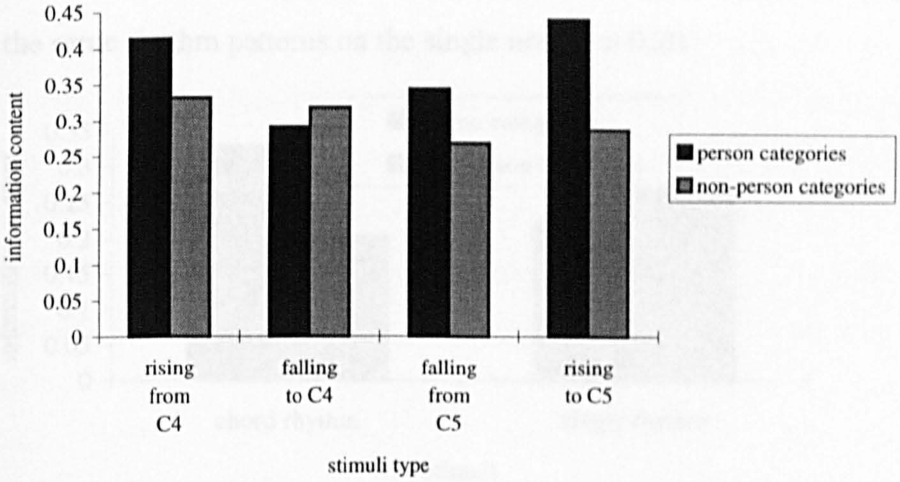


Figure 3 Person and non-person information in two-note intervals.
The y-axis displays the mean information content, whilst the x-axis shows the interval group. Responses are averaged across the twelve musical intervals in each of the four experimental groups.

A Pearson's chi-square statistic was used to determine whether response on the person categories was significantly different to response on the non-person categories. This was performed on proportion response counts. The two frequency counts from each category (person and non-person) are treated as observed frequencies in the chi-square analysis. The expected frequency for the analysis is the mean of the two observed counts. The two observed counts are tested against this expected count to determine whether they are significantly different.

Responses to the four person categories are compared with responses to the four non-person categories for all 48 musical intervals combined. The chi-squared analysis revealed that information recovered from the person categories is significantly higher than information recovered from the non-person categories in response to musical intervals, chi-squared (1 df) = 26.49, $p < 0.0001$.

The person - non-person comparison was also performed on responses to the repeating rhythm patterns. The information content in repeating rhythm patterns is shown below in Figure 4. Again, proportion response for all rhythm patterns is combined for the two different rhythm groups (C Major chord and single note, C4). The chi-squared analysis revealed that more information is recovered from the person categories than non-person categories for the rhythm patterns on the C Major chord, chi-squared (1 df) = 11.58, $p < 0.001$. There is no significant difference between person and non-person categories for the same rhythm patterns on the single note, $p > 0.20$.

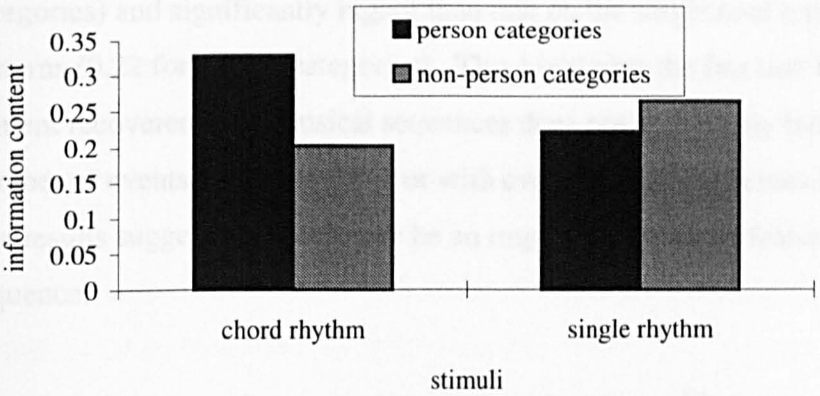


Figure 4 Information recovered from person and non-person categories in repeating rhythm patterns. The y-axis shows the mean information content whilst the x-axis displays the two rhythm groups. The different patterns were combined in each group.

A comparison of responses to two note intervals with responses to repeating rhythm patterns revealed that the average information recovered from person categories in two note intervals is significantly higher than information recovered from person categories in *single note* rhythm patterns, chi-squared (1 df) = 31.15, $p < 0.0001$. The person category response to *chord* rhythm patterns is not significantly different to that of the musical intervals, $p > 0.09$.

Discussion

The results of the present experiment (Experiment I), which examined responses to two note musical intervals, replicate the results of the original study which examined responses to large scale musical forms (Watt, 1995). As was found in the original experiment, the present results revealed that more information is recovered from person categories in comparison to non-person categories. The finding that information recovered in person categories is significantly different to information recovered in non-person categories for the two note intervals is an interesting one. This suggests that listeners are using the constructs typically used to deal with people in musical stimuli which are as short as two notes. Generally, the mean information content in the two note intervals is low: 0.38 and 0.29 for person and non-person categories, respectively.

This is low in comparison to the information content obtained in response to large scale musical forms which was 0.70 and 0.24 for person and non-person categories, respectively (Watt, 1995). However, the interval information content is similar to that obtained in the repeating rhythmic patterns on the C Major chord (0.33 for person categories) and significantly higher than that on the single note repeating rhythm patterns (0.22 for person categories). This highlights the fact that the information content recovered from musical sequences does not necessarily increase with increasing number of events in the sequence or with events containing a musical rhythm pattern. The results suggest that pitch may be an important structural feature of a musical sequence.

Very little is known about our perceptual responses to different musical intervals. There has been a long-standing agreement amongst researchers that listeners do not perceive musical intervals along one dimension alone (Helmholtz, 1954; Shepard, 1982; Krumhansl, 1990). Yet there is still no widely accepted account or explanation of what those other dimensions might be or what the best method of discovery might be. The present factor analysis of the interval data revealed that listeners perceive musical intervals along 3 main dimensions.

The first factor was described as interval size (the distance between the first and second tones of the pitch interval) with consonance. The general pattern of factor 1 scores shares some common features with results of studies which have examined the perceptual similarity of musical intervals. These previous studies have used methods such as rating the similarity of two successive tones on a dissimilar - similar scale (Allen, 1967; Kallman, 1982; Stoll & Parncutt, 1987; Parncutt, 1989) and less direct methods such as interval recognition using matching tasks and two alternative forced choice tasks between two melodic intervals (Balzano & Liesch, 1982; Parncutt, 1989; Stoll & Parncutt, 1987). The results of studies which examined similarity judgements of non-musicians to pure tones was that pitch interval size was the main determinant of similarity judgements (Kallman, 1982; Balzano & Liesch, 1982; Stoll & Parncutt, 1987; Parncutt, 1989). Perceived similarity of the unison interval was high, followed by a gradual decrease in similarity with increasing distance between the first and second tone of the interval.

When complex tones are used, the relationship between perceived similarity and pitch interval size is not monotonic. In many cases responses are best modelled by a tone chroma circle rather than a monotonic pitch height continuum (Deutsch, 1992; Balzano & Liesch, 1982). Parncutt (1989) found that similarity ratings were strongest for complex tones separated by an octave, perfect 4th and perfect 5th. Stoll & Parncutt (1987) also found that tones separated by an octave and perfect 5th were judged as being the most similar. It is argued that such findings reflect the perception of relationships between the harmonics in complex tones.

The experience of harmonic relationships is traditionally reported to result from the simultaneous presentation of tones. Intervals consisting of pitches at frequencies with small integer ratios, such as the octave (2:1), perfect 4th (4:3) and perfect 5th (3:2), result in the experience of consonance whilst intervals containing pitches with complex integer frequency ratios, such as the augmented 4th (45:32), result in the experience of dissonance (Burns & Ward, 1982). Stoll & Parncutt (1987) argue that the harmonic relationship between pitch intervals (which is strongest at the octave, then the perfect 4th and perfect 5th and so on) holds for simultaneous and successive intervals, although for different reasons.

As described in the methods section above (page 171), the tones used in the present study were complex. Factor 1 scores, which accounted for most of the variance in the data, showed a similar pattern to the results of perceived similarity ratings of successive pure (Kallman, 1982; Balzano & Liesch, 1982) and complex tones (Parncutt, 1989) with interval width effects and a differentiation between octave, perfect 4th and perfect 5th for intervals rising and falling to C_4 and rising to C_5.

It is not surprising that the present data set is comparable with results of more direct similarity studies given that psychological similarity/proximity is one of the most commonly used organisational principle in perceptual and cognitive domains (Krumhansl, 1983). Some tone pairs are psychologically more similar than others, with intervals close together generally perceived as more similar than intervals far apart (Krumhansl, 1990). Theories of pitch interval similarity have developed perceptual scales such as:

unison m2 M2 m3 M3 P4 A4 P5 m6 M6 m7 M7 octave
sameness.......similarity.......................differentiation........................

to explain responses to musical intervals (Narmour, 1992), which appear to be reflected in the pattern of factor 1 scores.

The second factor can be interpreted as small melodic intervals as distinct from large intervals. This would perhaps suggest an experience based differentiation of small versus large intervals. As discussed in Chapter 3, studies have shown that melodies, universally, use small pitch intervals between adjacent notes. Indeed there is a tendency to avoid intervals larger than 4 or 5 semitones in the majority of melodies (Dowling & Harwood, 1986). Statistical analyses have revealed that large intervals are relatively uncommon in melodies from a variety of genres ranging from classical - pop (Vos & Troost, 1989) and that minor and major seconds are the most common intervals to occur between successive tones (Shepard, 1982). The present pattern of results may reflect the tendency of listeners to differentiate between smaller more melodic intervals, which occur predominantly in much of the music we listen to, and larger intervals which are less common.

The present results were obtained using a non-direct 2AFC task. The principal components analysis yielded a similar pattern of results to previous experiments which have examined specifically the perceptual similarity of musical intervals (Balzano & Liesch, 1982; Parncutt, 1989). The results of the present study demonstrate that the 2AFC task is a useful tool with which to examine the perceptual responses of non-musical participants to basic musical stimuli. The information analysis showed that information content is less in the interval stimuli than the musical excerpts from the original experiment (Watt, 1995). Such a result is to be expected given that listeners are not accustomed to hearing isolated two note fragments. It is unlikely that such short two note fragments contain a high degree of meaningful information in isolation. The pitch intervals which listeners are more likely to experience are set within some kind of broader musical context. Experiment I has established that two note musical intervals, outwith a musical context, display the same person - non-person effect as large scale musical forms and convey similar amounts of information to repeating rhythm chord patterns. The information content increases somewhere between two note intervals and large scale musical forms. It is clear that this is not a simple effect of increasing notes such as those in a repeating rhythmical pattern. Pitch information is important. The interest now lies in finding out at which point the information recovered from a melody markedly increases. The same 2AFC technique was used with real Gaelic melodies in Experiment II to examine the perception of information content in unaccompanied Gaelic melody fragments of 2 - 8 notes in length. The main interest in this experiment lies in examining how the information content in a melody forms as the melody progresses.

Experiment II Perception of content in Gaelic melodies

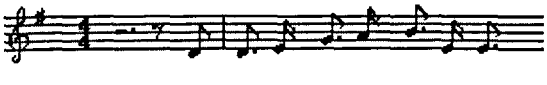
Method

Subjects The subjects were 183 first year psychology undergraduate students, 100 female and 83 male, aged between 17 and 35 years. The task was performed in first year laboratory sessions.

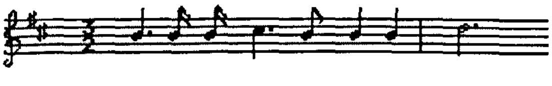
Stimuli Three melodies were chosen from a set which had been rated on a perceptual emotional quality scale (Gerstley, 1997, as discussed p.12). One melody was selected from each of the following perceived categories: happy, sad & neutral. These melodies are shown below in Figure B.

Figure B Melodies used in Experiment II

Happy (h)



Sad (s)



Neutral (n)

Extracts were either 2, 3, 4, 6, or 8 notes in length or 1.25 - 4.25 crotchet beats in length for melody h, 1.75 - 9 crotchet beats in length for melody s and 1.5 - 6 crotchet beats in length for melody n.

Procedure Listeners heard the melody fragment once and then performed the 2AFC task as above in Experiment I. They did this for *one* of the possible melody lengths from each of the three melodies and thus only ever heard *one* fragment from the same melody.

Analysis

As in Experiment I, the proportion response was counted separately for every musical stimulus. A Pearson's chi-squared was performed on the data for person and non-person categories to examine if there are significant differences between the information in these two categories. A second analysis of the information content in each of 2 - 8 note fragments for each melody was performed for person categories only. This was done to examine the points at which information content increased significantly in each melody. Information content in 2 note fragments were compared with 3 note fragments which were compared with 4 note fragments and so on. Thus the earliest point at which information significantly increases is identified in each melody.

Results

The average information content recovered from the person and non-person categories in response to the three melodies is displayed in Figures 6, 7 & 8, below. Mean information content is displayed separately for each melody as a function of length of melody to examine perceptual responses as the melody progresses.

The first comparison is between the person and non-person information in the three melodies. The chi-squared analysis, summarised in Table 1 below, revealed that there were significant differences between the information recovered from person and non-person categories which were dependent on length and type of melody. The person - non-person effect on information content occurs earliest in melody sad. There is a significant difference between the person and non-person categories in 3 note fragments of this melody with more information in the person categories in comparison to the non-person categories. This difference occurs in 4 note fragments for melody happy and not until 8 notes for melody neutral. Information is recovered more quickly via the person categories in the happy and sad melodies than it is in the neutral melody.

Table 1 Chi-Squared results of comparison between person and non-person categories.

Melody	2 notes	3 notes	4 notes	6 notes	8 notes
sad	---	4.44 +	11.00 *	15.31 *	---
neutral	---	---	---	---	13.77 *
happy	---	---	6.07 +	15.13 *	5.01 +

* p < 0.001 + p < 0.05 --- no significant differences p > 0.10 All df = 1

Of interest, is the point at which information content increases significantly as the melody progresses. The second set of comparisons are between the different lengths of each melody to ascertain where, in the melody, the information content increases. Person categories alone are used in this analysis. All three melodies show the same pattern of increased information content as the melody progresses. Generally, the highest information content is recovered from the complete phrase (8-note fragment) of each melody. (See Figures 6, 7 & 8, below).

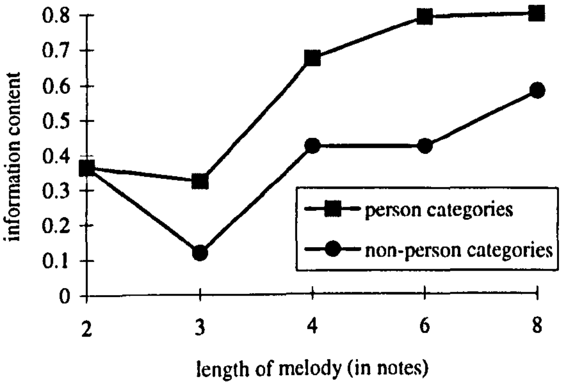


Figure 6 Information recovered through person compared with non-person categories for melody H. The y-axis displays the mean information content averaged across the four person and the four non-person adjective categories. Melody length is displayed in the x-axis.

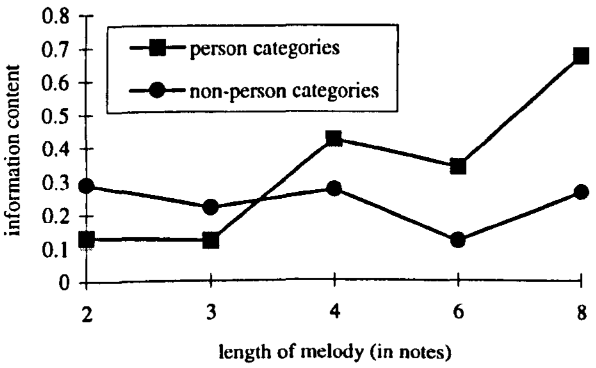


Figure 7 Information recovered through person compared with non-person categories for melody N. The y-axis displays the mean information content averaged across the four person and the four non-person adjective categories. Melody length is displayed in the x-axis.

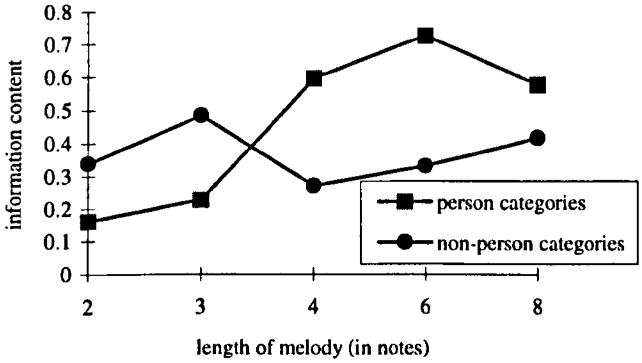


Figure 8 Information recovered through person compared with non-person categories for melody S. The y-axis displays the mean information content averaged across the four person and the four non-person adjective categories. Melody length is displayed in the x-axis.

The point at which there is a significant increase in information content occurs early on in each melody. For all three melodies, information in the 4 note fragment is significantly greater than information content in the 3 note fragments (see Table 2, below).

Table 2 Chi-Squared results of melody length comparison for person categories.

Melody	2 vs 3 notes	3 vs 4 notes	4 vs 6 notes	6 vs 8 notes
sad	---	10.57 **	---	---
neutral	---	6.52 +	---	9.83 **
happy	---	10.75 **	---	---

p < 0.001 ** p < 0.005 + p < 0.05 --- no significant differences p < 0.10 df = 1

Discussion

The present study shows a person - non-person effect which is similar to the effect obtained in the original study (Watt, 1995) and the interval study (Experiment I, above). The person - non-person difference does not occur in the first two note interval of the melodies as is the case for isolated intervals. This is best explained by the fact that the first two notes of each melody are on the same pitch. Experiment I revealed that information content in two note intervals is higher than in musical rhythms played on the same note. It is not, therefore, surprising that the musical intervals in Experiment I contain more information than the two note fragments of the melodies which are essentially two repeated pitches.

The findings of the present study also reveal that there are differences between each melody in the point at which information content recovered from person categories is significantly greater than information recovered from non-person categories. The person category difference is established in as little as 3 notes of the sad melody and 4 notes of the happy melody and not until the 8th note of the neutral melody. This offers evidence for an effect of melody mood on information recovered from person and non-person categories. These results suggest that melodies which are rated as more emotionally extreme (sad and happy melodies) evoke person like attributes earlier on than those which are rated as emotionally neutral.

The length comparison for each melody yielded similar results for the three different melodies. There is significantly more information content in 4 note fragments of the melodies than in 3 note fragments. There are no significant increases in information content in the shorter fragments 2 - 3 note fragments. The increase in information at 4 notes is the first significant increase in information to occur in all the melodies. This suggests that there is something interesting happening to the perceived information content after 4 notes of these melodies. It is interesting that listeners can establish a firm underlying beat to Gaelic melodies after hearing as little as 5 - 6 notes (Chapter 3). The finding that beat information is communicated sufficiently at around 5 - 6 notes ties in well with the present finding that information content increases at 4 notes of the melodies. The earlier occurrence of the information content increase revealed by the present experiment may be due to the fact that the information examined is not limited to beat information as it is in the beat perception study (Chapter 3).

This non-musical task is clearly useful in uncovering the listeners' sensitivity to information units in the Gaelic melodies. Given that such effects occur in these melodies, it is also of interest to examine the relative contributions of pitch and rhythm factors to information content. The 2AFC technique is used in Experiment III to establish the role of rhythm and pitch factors separately as the melody progresses.

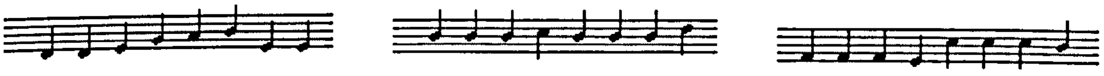
Experiment III Perception of content in structurally altered Gaelic melodies

The same 2AFC task was used in response to stimuli which contained the same pitches or rhythms as melodies used in Experiment II. Performance on the task in response to the original melodies (Experiment II) is compared with response to pitches alone and rhythms alone to examine the contribution of these factors to perceptual responses.

Subjects The subjects were 77 psychology undergraduate students, 46 females and 31 males, aged between 17 and 38 years.

Stimuli The pitch and rhythm patterns of the three melodies used in Experiment II were utilised in the present experiment. The pitches were played equitemporally (every note was a crotchet) and the rhythms were played on one note alone (the note on which each melody began). These excerpts were either short: the first four notes of the fragments in Figure C (below) or long: the complete eight note sequence.

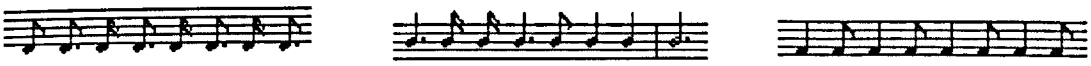
Figure C The pitch and rhythm patterns used in Experiment III



Happy pitch pattern

Sad pitch pattern

Neutral pitch pattern



Happy rhythm

Sad rhythm

Neutral rhythm

Again, the sequences were played without accent at a tempo of one crotchet beat per 0.60 seconds.

Procedure

Listeners heard the melody fragment once and then performed the 2AFC task as above in Experiment I. They did this for *one* of the excerpts from each melody and thus only ever heard *one* fragment from the same melody.

Analysis

A measurement of *information content* was calculated in the same way as Experiments I and II for person categories. Comparisons were made between pitch alone, rhythm alone and original melody condition using a chi-squared analysis.

Results

The mean information content results for each condition in each melody are shown in Figures 9, 10 & 11, below. The results of the chi-squared analysis are shown in Tables 3 & 4, below.

Melody h

In response to melody h, there is a significant difference between melody and rhythm alone conditions with more information in the melody. This effect occurs in short and long fragments. There is no statistical difference between melody and pitch alone or pitch alone and rhythm alone conditions.

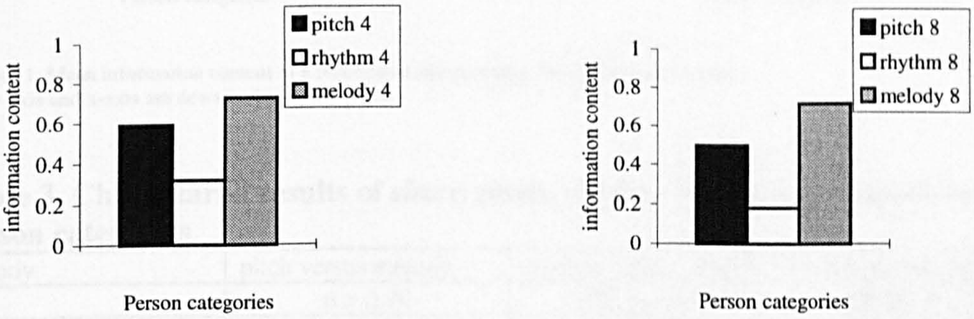


Figure 9 Mean information content as a function of category type for happy melody excerpts.
The y-axis displays mean information content which is averaged across person adjective categories. The x-axis shows the stimuli: short 4 note and long 8 note excerpts of pitch alone, rhythm alone and original melody.

Melody n

There are no statistical differences between any of the conditions in the short fragment of melody n. There is a significant difference between melody and rhythm alone in the long fragment with more information in the original melody.

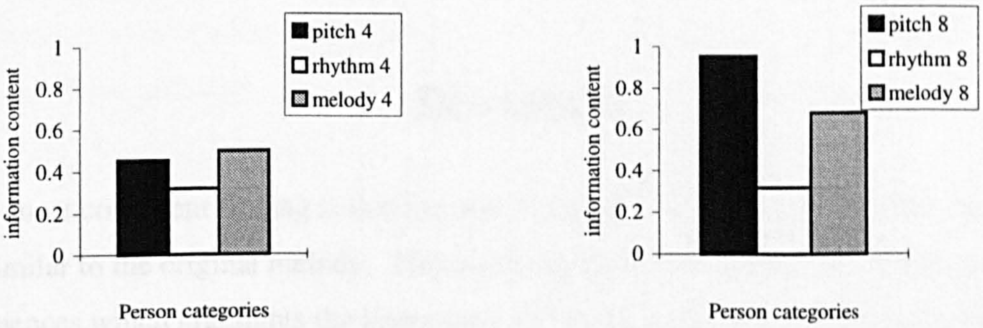


Figure 10 Mean information content as a function of category type for neutral melody excerpts.
The y-axis and x-axis are described in detail above.

Melody S

For the short fragment of melody s, the pitch alone structure and melody conditions convey significantly more information than the rhythm alone condition. These effects disappear in the longer 8-note fragment where there is a trend for increased information in the rhythm alone condition.

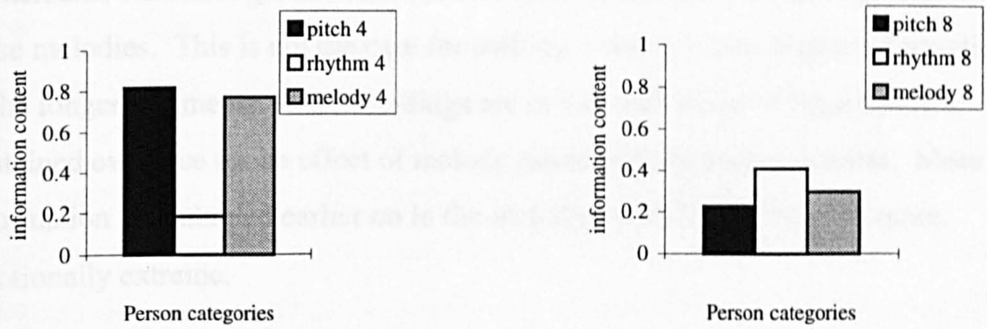


Figure 11 Mean information content as a function of category type for sad melody excerpts. The y-axis and x-axis are described in detail above.

Table 3 Chi-Squared results of *short*: pitch, rhythm & melody comparison of person categories.

Melody	pitch versus melody	rhythm versus melody	pitch versus rhythm
sad	$p > 0.70$	8.48, $p < 0.005$	10.24, $p < 0.005$
neutral	$p > 0.80$	$p > 0.20$	$p > 0.20$
happy	$p > 0.50$	4.00, $p < 0.05$	$p > 0.20$

All df = 1

Table 4 Chi-Squared results of *long*: pitch, rhythm & melody comparison of person categories.

Melody	pitch versus melody	rhythm versus melody	pitch versus rhythm
sad	$p > 0.90$	$p > 0.70$	$p > 0.40$
neutral	$p > 0.06$	$p > 0.08$	12.91, $p < 0.001$
happy	$p > 0.30$	5.80, $p < 0.005$	$p > 0.10$

All df = 1

Discussion

The most consistent finding is that the amount of information in pitch alone sequences is similar to the original melody. This result occurs for all melodies in short and long sequences which highlights the importance of pitch information in these melodies. The remaining results are specific to each melody. The main finding is that where significant effects occur, the rhythm alone conveys less information than either melody or pitch alone conditions. Melody h contains more information than rhythm alone conditions for both short and long sequences. Melody s and the pitch alone sequence contain more information than the rhythm alone for the short condition only. There is more information in the pitch condition than rhythm condition in the long fragment of melody n. It is interesting to note that the information contained in the short fragments of melodies s and h is greater than the information contained in the long fragments of these melodies. This is not the case for melody n which yields higher information levels in the longer fragments. These findings are in line with those of Experiment II which contained evidence for an effect of melody mood on information content. More information is contained earlier on in the melodies which were rated as more emotionally extreme.

Summary

The present set of experiments reveal that the 2AFC task is a useful non-musical task with which to examine the listeners' knowledge and perception of musical structure in basic musical stimuli, such as musical intervals, and in unaccompanied melodies. When used with basic isolated two note musical intervals, this task reveals a sensitivity to interval size and harmonic relationships (which have been revealed by previous studies using more direct methods) and small conjunct versus large disjunct pitch intervals. Using this technique with excerpts from Gaelic melodies reveals effects of length and mood of the melody on the information which listeners recover and the possible meaning they perceive. The present results suggest that listeners can extract a basic meaning via person categories from as little as 3 and 4 notes of a melody (melodies s and h) and from two note musical intervals. In addition to this, the findings suggest that the amount of information communicated in these melodies increases once listeners hear 4 notes of the melodies.

Chapter 6 Discussion

The present thesis reports on a series of tasks which were developed to examine the perception of structure, at a number of different levels, in musical sequences such as Gaelic melodies and musical intervals. Recall that the basic musical structures which were studied in this thesis were phrase structure, rhythm and pitch and musical beat. These were outlined briefly in Chapter 1 and focused on individually in each of the following chapters. The thesis was organised into four main experimental chapters which were ordered such that research on the larger perceptual structures, i.e. phrase, was presented first and was followed by work which examined progressively shorter units of musical structure such as the beat and two note musical intervals. Work which examined the position and size of the units into which listeners segment and process melodies was presented in the first experimental chapter and this was followed by research on the memorable rhythmic unit. Perceptual responses at the level of the musical beat were then dealt with and, finally, responses to short two note melodic musical intervals and fragments of melodies were presented. The main findings of interest from each experimental chapter will be discussed in turn and then followed by a summary of what the results collectively tell us about the way in which listeners perceive musical information.

Summary

One chunk or two?

The segmentation and click localisation research examined the position at which listeners segment Gaelic melodies and investigated how listeners process information in the melodies in relation to perceived points of segmentation. The number of chunks into which listeners segmented the melodies was not of prime importance. Rather, it was the size of the chunks and the features of the musical stimulus which elicited the segmentation response which were of interest. The results revealed that consistent segmentation occurred in melodies which contained salient boundary markers, such as long notes, and also highlighted the fact that most listeners did not choose to segment the melodies into units which were longer than 5.4 seconds or shorter than 2.4 seconds in duration.

As discussed in the introduction of Experiment II, Chapter 2, the present click localisation technique was modified to cater for a participant sample of non-musicians and to eliminate previous methodological problems, such as the ear effect prevalent in the dichotic listening technique and the confounding factor of visual input from the musical score. The results of the click localisation task revealed that listeners process information as discrete units which correspond to the units into which listeners perceptually segment the melodies. Clicks within each segmented unit were migrated towards the nearest perceptual boundary which, in line with previous click localisation studies with musicians, suggests that listeners attempted to preserve the coherence of phrasal groups.

It became apparent in the click localisation experiment that listeners could not accurately process much more than two chunks (or segments) of information in a melody in the detail which was required for the localisation task. This was shown by the poor performance of click localisation of click positions which were near the end of a three phrase excerpt in one of the melodies. This emphasised that there are temporal constraints which must be adhered to when using the modified localisation task. These constraints are due mainly to the demands of rehearsing large amounts of musical information in memory whilst waiting for the click position to occur in the repeated melody extract. If, however, the click localisation technique is used sensibly, and musical excerpts are kept within a ten second duration, then it can be used successfully to examine the way in which listeners process auditory information, without the visual input from a musical score.

How much can you remember?

The rhythm recall experiment examined how much rhythmic information listeners were capable of remembering after listening to an unfamiliar melody. This revealed that listeners could recall an average amount of rhythmic information which was equivalent to 2.4 - 3 seconds of the Gaelic melodies. The amount of rhythm information recalled from the Gaelic melodies was generally shorter than the perceived phrasal units which were identified by the Segmentation Experiment (Chapter 2).

The importance of meaningful and repetitive units in the musical rhythms in facilitating recall was highlighted by the finding that listeners recalled less of the computer generated melodies (1.8 seconds) that they recalled of the Gaelic melodies (2.4 - 3 seconds). As discussed in Chapter 3, the Gaelic melodies consisted of a series of repetitive rhythmic units whereas the durations in the computer generated melodies were created by a series of probabilistic rules which did not result in the repetitive rhythmic units prevalent in the Gaelic melodies. The other result which supports the beneficial effect of repetitive rhythmic units upon recall was that the amount of rhythm information which listeners recalled from computer generated sequences was a function of the length of the sequence. Rhythm recall decreased as the sequence length increased. In contrast to this, the amount of rhythmical information recalled from real melodies was not a simple function of sequence length.

A number of experimental conditions compared the effects of pitch structure on rhythm recall of Gaelic and computer generated melodies. Recall of unstructured and structured sequences were compared. The structured sequences were tonal (in the key of C major) and the unstructured sequences were atonal (constructed from all the notes in the octaves immediately above and below C_4). The structured patterns also contained small pitch intervals (1 - 4 semitones) whilst the unstructured patterns contained predominantly large pitch intervals (greater than 4 semitones and ranging up to 2 octaves). The results revealed that recall of computer generated rhythms was better for structured sequences than for unstructured sequences. When the pitch patterns of Gaelic melodies were structurally altered by reversal there was no effect on rhythm recall. The differences between the pitch intervals in the original Gaelic melody versus reversed pitch Gaelic melody conditions of Experiment II were not as extreme as the differences between structured and unstructured pitch patterns of Experiment I. In the unstructured versus structured sequences there were differences of tonality as well as pitch interval size which may have resulted in poorer recall of these sequences.

The reversed pitch patterns of the Gaelic melodies used in Experiment II were in the same key as the original melody and contained the same size of intervals as the original (mainly 0 - 3 semitones) but in the opposite direction (i.e. ascending and descending intervals in the original melody became descending and ascending intervals in the reversed melody, respectively). The reversed pitch condition did not, however, contain the meaningful combination of pitch and rhythm structures which the original melody contained and it was expected that this may have an effect on rhythm recall. Pitch reversal, which was less extreme than pitch alteration between unstructured and structured sequences, did not, however, effect rhythm recall of the Gaelic melodies used in the present study.

Can you feel the beat?

A beat-tapping task was used to examine the perception of beat in melodies. This simple technique has been used widely in previous studies (Handel & Oshinsky, 1983; Vos et al, 1994). Unlike previous beat-tapping studies, the main interest in this experiment lay in examining how soon into the melody listeners could perceive the beat and how information about the beat structure was communicated as the melody progressed. This experiment revealed that the majority of listeners could perceive an underlying beat in as little as 5 or 6 notes or 1.5 - 3 seconds worth of information of the melodies. In addition to this, most information about the beat was communicated within the first phrase of each melody as was shown by steadily decreasing standard deviations as the melody progressed. This effect was not mirrored in the second phrase which, in line with theoretical work on perception of metre in music (Steedman, 1977; Longuet-Higgins & Lee, 1982; Lee, 1991), emphasises the importance of information at the beginning of a melody in establishing the listeners beat perception.

Non-traditional technique

The 2AFC experiment was used to examine the perception of structure and content in musical intervals and fragments of Gaelic melodies in a constrained and non-biased manner. Listeners performed the 2AFC task in response to the musical stimuli. The main difference between the present task in comparison to previous tasks which have attempted to examine meaning in music is that this task was non-direct. There was no mention of meaning, content or emotion and no instructions requiring listeners to focus on a particular aspect of the musical structure. The results of responses to musical intervals showed evidence of perceptual similarity of the pitch intervals based mainly on interval size. Sensitivity to similarities between tones in successive pitch intervals has been demonstrated previously using direct methods, such as similarity rating tasks

(Kallman, 1982; Parncutt, 1989; Balzano & Liesch, 1982; Stoll & Parncutt, 1987) and less direct methods such as interval matching tasks and two alternative forced choice tasks (Balzano & Liesch, 1982; Parncutt, 1989; Stoll & Parncutt, 1987).

The results of the present 2AFC experiment also demonstrated that the point at which information content significantly increased in the three melodies occurred between 3 and 4 notes or 2.25 - 3.5 crotchet beats (1.35 - 2.1 seconds). In addition to this, the person - non-person category effect which was found in response to large scale musical forms and musical intervals occurred earlier in the more emotionally charged melodies than it did in the neutral melody (at 3 and 4 notes for the sad and happy melodies compared with 8 notes in the neutral melody).

Conclusions

The common theme of all five experimental tasks is that they examined the listeners' perceptual responses to structural features of the musical stimulus. All the tasks were developed so that they could be utilised with a participant sample of non-musicians and were designed so as to have a minimally disruptive effect on the listening process. The tasks were novel, (2AFC task), or were a novel modification (click localisation task) of previous techniques which had been designed for use with musicians (Gregory, 1978; Sloboda & Gregory, 1978; Stoffer, 1985). The other three experimental tasks (segmentation, rhythm reproduction and beat-tapping) have been used in previous research (Krumhansl, 1996; Drake, 1993; Vos et al, 1994). However, in previous studies the tasks were utilised for purposes which were different to those of the present research and used materials which were unlike the unaccompanied melodies examined in the present set of studies. So in this respect, the tasks were either novel or examined a novel aspect of the perceptual response.

The 2AFC task was the most non-musical task of those used in the thesis; the experiment did not examine one particular aspect of the musical structure in the way that the other four tasks did. The particular structural feature of the music which was examined by the click localisation and segmentation tasks was not explicitly mentioned. Phrases were described generally in an unconstrained way as units of meaning. Listeners were not directed towards a particular structural feature of the music when performing the segmentation, click localisation and 2AFC tasks in the way that they were in the beat-tapping and rhythm recall tasks. Listeners were requested to respond to the beat and the rhythm of the music in the beat-tapping and rhythm recall tasks, respectively. Despite the fact that these tasks directed attention towards a particular feature of the music, instructions were kept as general and non-specific as possible. In the beat-tapping task, no mention was made of the metrical level at which listeners should respond. The rhythm recall task was described as a rhythm reproduction task and listeners were informed that their ability at tapping a rhythm was being assessed rather than their STM for the rhythmic sequences.

Despite using a number of different experimental tasks, the results can be interpreted alongside each other in a complementary way. The units into which listeners segmented the Gaelic melodies were found to correspond to results obtained in other experimental tasks such as beat perception and click migration which, as noted above, did not explicitly direct the listeners' attention towards the phrase structure of the melodies. The click localisation task resulted in the perceptual migration of superimposed clicks in the melodies towards the main phrase boundaries. The results of the beat-tapping task showed that information increased steadily up until the end of the first perceived phrase of the melodies. The information units identified in the segmentation task thus corresponded to some kind of perceptual unit which was uncovered by the click localisation and beat-tapping task.

The rhythm recall experiment revealed that listeners recalled an amount of rhythmic information which was shorter than the units into which listeners consistently segmented the melodies. The memorable rhythmic unit was similar in length to the amount of melody which was required to establish a strong sense of the underlying beat, as was demonstrated by the beat perception task. The rhythm recall and beat-tapping tasks both revealed that repetitive units were of the utmost importance in facilitating memory for and beat perception of the Gaelic melodies. As is suggested by Ockelford (1991), repetition is the one universal characteristic of musical patterning which contributes to the listeners' understanding of the musical structure. The present results demonstrate the way in which repetition can enable the listener to perceive and remember the rhythmic structure of the music.

The average memorable rhythmic unit and length of melody required for a consistent beat to be established was longer than the length of melody needed for a basic meaning to be perceived, as was shown in the 2AFC task. It is interesting that there was an effect of melody mood recovered by the 2AFC task. The point at which person information is significantly higher than non-person information occurred earliest in the emotionally extreme melodies. There was also an effect of melody mood on results of the beat perception experiment. The minim beat was established in 1.5 seconds in the melody perceived as being most happy and not until 2.4 seconds in the melody rated as being the most sad. Furthermore, the rate at which listeners tapped to the beat was faster in the happy melody than in the sad melody. These perceptual responses may be directly related to the communicative function of each melody.

Taken together, these results suggest that listeners process the melodies into units of information which they perceptually differentiate from each other. The rhythm recall experiment demonstrated that the average length of these units exceeds STM capacity. So although these perceptual groups may be used to chunk the melodies into shorter, more memorable units, listeners do not remember the specific details of these units. The fact that listeners remember an amount of rhythmic information which is greater than the amount of melody required to establish a basic meaning may be due to the memory facilitation of repetitive units within the musical structure. It is noteworthy that the amount of melody required to establish a perceptual group or the musical beat and the amount of rhythmic structure which is remembered all incorporate the minimal amount of information which is required for a basic meaning.

There are limitations in the present research. The perceptual responses were obtained using a limited number of Gaelic melodies which limits the conclusions which can be drawn. However, the present experimental techniques could easily be used with other types of music. Most music will contain perceptual segmentation points which could reflect perceptual information groups which can be tested using the tasks developed in the present study. Segmentation responses to other kinds of music, without the salient boundary markers which characterised many of the Gaelic melodies, may not be as consistent as those obtained in response to the Gaelic melodies. Listeners will, nonetheless, perceptually segment music they hear into smaller units, perhaps with a larger individual difference and less consistency as the segmentation responses to the Gaelic melodies under examination in the present thesis.

In addition to this, the present set of tasks could be extended for use with music which has been expressively performed by musicians rather than on computer to establish what effect, if any, expressive performance deviations may have on the way in which listeners perceive structure in music. It would be of interest to examine perceptual responses to a wider range of music using the tasks which were developed in the present study.
Using these tasks, experiments which manipulate direction and focus of attention to specific features of the musical structure could be carried out to examine the way in which perceptual responses depend on the particular feature of the musical structure which is attended to (Boltz, 1997). Attention could be directed towards, for example, the pitch of the musical excerpts to examine the effect on performance of beat-tapping, rhythm recall, segmentation or click localisation tasks. Comparisons could be made with performance of these tasks when attention is focused on, for example, the rhythmic structure.

Although there were no effects of familiarity with the Gaelic melodies on click localisation performance in the present thesis, the performance of unfamiliar versus familiar participants to the present tasks could be compared using a wider range of music. As discussed in Chapter 2, the lack of familiarity effect in the present click localisation experiment may have been due to the particularly salient boundary markers in the Gaelic melodies. There is no reason to suppose that there will be no effect of familiarity with other types of music. In addition to this, performance on tasks before and follow-up studies of performance after repeated listening to different types of music may provide additional information about how perceptual groups form from smaller into larger units of information. Given that the tasks used in this thesis are relatively simple to understand and perform, their use could also be extended to younger participant samples to study the way in which music is processed from a developmental aspect.

The present set of results imply that tasks which can be easily understood and performed by listeners without a musical background can uncover interesting features about the way in which listeners perceive auditory information. The present research has demonstrated that the tasks developed and used in the present thesis are a useful set of experimental techniques with which to examine the way in which listeners perceive structure in basic musical fragments and Gaelic melodies. It hoped that these will be employed to examine perceptual responses to other types of music.

References

Aiello, R. (1994). Music and language: parallels and contrasts. In *Musical Perspectives*, eds., R. Aiello and J. A. Sloboda, Oxford University Press, pp 40 - 63.

Allen, D. (1967). Octave discriminability of musical and non-musical subjects. *Psychonomic Science*, 7, 421 - 422.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In *The Psychology of Learning and Motivation*, Vol 2, eds., K. W. Spence, J. T. Spence, Academic Press, New York, pp 89 - 195.

Attneave, F., & Olson, R. K. (1971). Pitch as medium: a new approach to psychophysical scaling, *American Journal of Psychology*, 84, 147 - 166.

Baddeley, A. D. (1990). *Human Memory: Theory and Practice*. Lawrence Erlbaum Associates, London.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In *The Psychology of Learning and Motivation*, Vol 8, ed. G. Bower, Academic Press, San Diego, California, pp 47 - 89.

Balaban, M., Ebcioglu, K., & Laske, O. (1992). *Understanding Music with AI: Perspectives on Music Cognition.* AAAI Press/MIT Press, Cambridge.

Balzano, G. J., & Liesch, B. W. (1982). The role of chroma and scalestep in the recognition of musical intervals in and out of context. *Psychomusicology*, 2, 3 - 31.

Benjamin, W. E. (1984). A theory of musical meter. *Music Perception*, 1, 355 - 413.

Bigand, E. (1993). Contributions of music to research on human auditory cognition. In *Thinking in Sound: The Cognitive Psychology of Human Audition*, eds., S. McAdams and E. Bigand, Oxford University Press, pp 231 - 277.

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

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

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

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

Boltz, M. G. (1997). The processing of melodic and temporal information: independent or unified dimensions? In *Proceedings of the Second Meeting of the FWO Research Society on Foundations of Music Research "Foundations of Time Perception"*, eds., D. Moelants and M. Leman, University of Ghent, pp 17 - 27.

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

Boring, E. G. (1950). *A History of Experimental Psychology.* 2nd edition New York: Appleton-Century-Crofts, 1950.

Bower, G. H., & Winzenz, D. (1969). Group structure, coding, and memory for digit series. *Journal of Experimental Psychology, Monograph*, 80, 1 - 17.

Boyden, D. D. (1971). *An Introduction to Music.* 2nd edition, Faber and Faber Ltd., London.

Brandt, W., Corra, A., Christ, W., Delone, R., & Winold, A. (1980). *The Comprehensive Study of Music Vol VI: Basic Principles of Music Theory.* Harper & Row Publishers, New York, London.

Bregman & Campbell (1971). Primary auditory stream segregation and perception of order in rapid sequences of tones. *Journal of Experimental Psychology*, 89, 244 - 249.

Burns, E. M., and Ward, D. W. (1978). Categorical perception - phenomenon or epiphenomenon: evidence from experiments in the perception of melodic musical intervals. *Journal of the Acoustical Society of America,* 63, 456 - 468.

Burns, E. M., and Ward, D. W. (1982). Intervals, scales and tuning. In *The Psychology of Music*, ed. D. Deutsch, Academic Press, London, pp 241 - 269.

Clarke, E. F. (1982). Timing in the performance of Erik Satie's 'vexations', *Acta Psychologica*, 50, 1 - 19.

Clarke, E. F. (1985). Structure and expression in rhythmic performance. In *Musical Structure and Cognition*, eds., P. Howell, I. Cross and R. West, Academic Press, London, pp 209 - 236.

Clarke, E. F. (1987). Categorical Rhythm Perception: An ecological perspective. In *Action and Perception in Rhythm and Music*, ed., A. Gabrielsson, Royal Swedish Academy of Music, pp 19 - 33.

Clarke, E. F., & Krumhansl, C. L. (1990) Perceiving musical time. *Music Perception*, 7, 213 - 252.

Cook, N. (1994). Perception: a perspective from music theory. In *Musical Perspectives*, eds., R. Aiello and J. A. Sloboda, Oxford University Press, pp 64 - 95.

Cross, I., Howell, P., and West, R. (1983). Preferences for scale structure in melodic sequences. *Journal of Experimental Psychology: Human Perception and Performance*, 9, 444 - 460.

Crowder, R. G. (1993). Auditory Memory. In *Thinking in Sound: The Cognitive Psychology of Human Audition*, eds., S. McAdams and E. Bigand, Oxford University Press, pp 113 - 145.

Crozier, J. B. (1974). Verbal and exploratory responses to sound sequences varying in uncertainty level. In *Studies in the New Experimental Aesthetics*, ed., B. E. Berlyne, pp 27 - 90.

Cuddy, L. L., Cohen, A. J., & Mewhort, D. J. K. (1981). Perception of structure in short melodic sequences. *Journal of Experimental Psychology: Human Perception & Performance*, 7, 869 - 883.

Dalrymple-Alford, E. C. (1976). Response bias and judgement of the location of clicks in sentences. *Perception & Psychophysics*, 19, 303 - 308.

Davies, J. B. (1978). Tonality's poor cousin. In *The Psychology of Music*, pp 176 - 200.

Deliege, I. (1987). Grouping conditions in listening to music: an approach to Lerdahl & Jackendoff's grouping preference rules. *Music Perception*, 4, 325 - 360.

Desain, P., & Honing, H. (1997). Computational models of beat induction: the rule based approach. In *Proceedings of the Second Meeting of the FWO Research Society on Foundations of Music Research "Foundations of Time Perception"*, eds., D. Moelants and M Leman, University of Ghent, pp 34 - 43.

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

Deutsch, D. (1982a). *The Psychology of Music*. Ed., D. Deutsch, Academic Press, San Diego, California.

Deutsch, D. (1982b). Grouping mechanisms in music. In *The Psychology of Music*, ed., D. Deutsch, Academic Press, San Diego, California, pp 99 - 134.

Deutsch, D. (1983). The generation of two isochronous sequences in parallel. *Perception & Psychophysics*, 34, 331 - 337.

Deutsch, D. (1992). The tritone paradox: implications for the representation and communication of pitch structures. In *Cognitive Bases of Musical Communication*, eds., M. R. Jones and S. Holleran, American Psychological Association, Washington, DC, pp115 - 138.

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

Dowling, W. J. (1973). 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. (1991). Tonal strength and melody recognition after long and short delays. *Perception & Psychophysics*, 50, 305 - 313.

Dowling, W. J. & Bartlett, J. C. (1981). The importance of interval information in long-term memory for melodies. *Psychomusicology*, Spring, 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.

Dowling, W. J., & Harwood, D. L. (1986). *Music Cognition.* Academic Press, London.

Drake, C. (1993). Reproduction of musical rhythms by children, adult musicians, and adult nonmusicians. *Perception & Psychophysics*, 53, 25 - 33.

Drake, C., & Palmer, C. (1993). Accent structures in music performance. *Music Perception*, 10, 343 - 378.

Drake, C., Penel, A., Bigand, E., & Stefan, L. (1997). Tapping in time with musical and mechanical sequences. In *Proceedings of the Third Triennial ESCOM Conference*, ed., A. Gabrielsson, Uppsala University.

Edworthy, J. (1985). Melodic contour and musical structure. In *Musical Structure and Cognition*, eds., P. Howell, I. Cross and R. West, Academic Press, London, pp 169 - 188.

Fisher, R. A. (1950). *Contributions to Mathematical Statistics.* John Wiley & Sons, Inc, Canada.

Fitzgibbons, D. J., Pollatsek, A., & Thomas, I. B. (1974). Detection of temporal gaps within and between perceptual tonal groups. *Perception & Psychophysics*, 16, 522 - 528.

Fodor, J. A., & Bever, T. G. (1965). The psychological reality of linguistic segments. *Journal of Verbal Learning and Verbal Behaviour*, 4, 414 - 420.

Fodor, J. A., Bever, T. G., & Garrett, M. F. (1974). *The Psychology of Language.* New York: McGraw-Hill.

Fraisse, P. (1964). *The Psychology of Time.* (trans. J. Leith Ellis from 1st edition 1963), Eyre & Spottiswoode, Ltd.

Fraisse, P. (1982). Rhythm and tempo. In *The Psychology of Music*, ed. D. Deutsch, Academic Press, London, pp149 - 180.

Fraisse, (1987). A historical approach to rhythm as perception. In *Action and Perception in Rhythm and Music*, ed., A. Gabrielsson, Royal Swedish Academy of Music, pp 7 - 19.

Frankish, C. (1989). Perceptual organisation and PAS. *Journal of Experimental Psychology: Human Learning & Memory*, 15, 469 - 479.

Frankish, C. (1995). Intonation and auditory grouping in immediate serial recall. *Applied Cognitive Psychology*, 9, 5 - 22.

Gabrielsson, A. (1973). Adjective ratings and dimension analyses of auditory rhythm patterns. II. *Scandinavian Journal of Psychology*, 244 - 260.

Gabrielsson, A. (1974). Performance of rhythm patterns. *Scandinavian Journal of Psychology*, 15, 63 - 72.

Gabrielsson, A., & Juslin, P. N. (1996). Emotional expression in music performance: between the performer's intention and the listener's experience. *Psychology of Music*, 24, 68 - 91.

Gerard, C., & Auxiette, C. (1988). The role of melodic and verbal organisation in the reproduction of rhythmic groups by children. *Music Perception*, 6, 173 - 192.

Gerstley, L. D. (1997). Cross-modal and synaesthetic perception in music and vision. *Unpublished PhD thesis*, Stirling university, Stirling.

Gfeller, K., Asmus, A., & Eckert, M. (1991). An investigation of emotional responses to music and text. *Psychology of Music*, 19, 128 - 141.

Gregory, A. H. (1978). Perception of clicks in music. *Perception & Psychophysics*, 24, 171 - 174.

Handel, S. (1984). Using polyrhythms to study rhythm. *Music Perception*, 1, 465 - 484.

Handel, S., & Lawson, G. R. (1983). The contextual nature of rhythmic interpretation. *Perception & Psychophysics*, 34 (2) 103 - 120.

Handel, S., & Oshinsky, J. S. (1981). The meter of syncopated auditory polyrhythms. *Perception & Psychophysics*, 30 (1), 1 - 9.

Helmholtz, H. (1954). *On the sensations of tone as a physiological basis for the theory of music* (trans. A. J Ellis from 4th edition 1877), Dover, New York (reprinted from English 2nd edition)

Hevner, K. (1936). Experimental studies of the elements of expression in music. *American Journal of Psychology*, 48, 246 - 268.

Hulme, C., & Tordoff, V. (1989). Working memory development: the effects of speech rate, word length and acoustic similarity on serial recall. *Journal of Experimental Child Psychology*, 47, 72 - 87.

Huttenlocher, J., & Burke, D. (1969). Why does memory span increase with age? *Cognitive Psychology*, 8, 1 - 31.

Jackendoff, R. (1991). Musical parsing and musical affect. *Music Perception*, 9, 199 - 230.

Johnson, N. F. (1970). The role of chunking and organisation in the process of recall. In G. H. Bower (Ed.), *Psychology of learning and motivation*. New York: Academic Press.

Jones, M. R. (1976). Time, our lost dimension: toward a new theory of perception, attention, and memory. *Psychological Review*, 83, 323 - 355.

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, 211 - 218.

Jones, M. R. & Holleran, S. (1992). Cognitive bases of musical communication: an overview. In *Cognitive Bases of Musical Communication*, eds., M. R. Jones and S. Holleran, American Psychological Association, Washington, DC, pp 1 - 6.

Jones, M. R., Maser, D. J., & Kidd, G. R. (1978). Rate and structure in memory for auditory patterns, *Memory & Cognition*, 6, 246 - 258.

Jones, M. R., & Yee, W. (1993). Attending to auditory events: the role of temporal organisation. In *Thinking in Sound: The Cognitive Psychology of Human Audition*, eds., S. McAdams and E. Bigand, Oxford University Press, pp 69 - 112.

Jusczyk, P. W., & Krumhansl, C. L. (1993). Pitch and rhythmic patterns affecting infants' sensitivity to musical phrase structure. *Journal of Experimental Psychology: Human Perception & Performance*, 19, 627 - 640.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, 20, 141 - 151.

Kallman, H. J. (1982). Octave equivalence as measured by similarity ratings. *Perception & Psychophysics*, 32, 37 - 49.

Kelso, J. A. S., Holt, K. G., Rubin, P., & Kugler, P. N. (1981). Patterns of human interlimb coordination emerge from the properties of non-linear, limit cycle oscillatory processes: theory and data. *Journal of Motor Behavior*, 13, 226 - 261.

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

Klapp, S. T., Hill, M. D., Tyler, J. G., Martin, Z. E., Jagacinski, R. J., & Jones, M. R. (1985). On marching to two different drummers: perceptual aspects of the difficulties. *Journal of Experimental Psychology: Human Perception & Performance*, 11, 814 - 827.

Krumhansl, C. L. (1983). Perceptual structures for tonal music. *Music Perception*, 1, 28 - 62.

Krumhansl, C. L. (1990). *Cognitive Foundations of Musical Pitch.* Oxford University Press, New York, Oxford.

Krumhansl, C. L. (1996). A perceptual analysis of Mozart's piano sonata K.282: segmentation, tension and musical ideas. *Music Perception*, 13, 401- 432.

Krumhansl, C. L., & Jusczyk, P. W. (1990). Infants perception of phrase structure in music. *Psychological Science*, 1, 70 - 73.

Ladefoged, P., & Broadbent, D. E. (1960). Perception of sequence in auditory events. *Quarterly Journal of Experimental Psychology*, 12, 162 - 170.

Lee, C. S. (1985). The rhythmic interpretation of simple musical sequences: towards a perceptual model. In *Musical Structure and Cognition*, eds., P. Howell, I. Cross and R. West, Academic Press, London, pp 53 -69.

Lee, C. S. (1991). The perception of metrical structure: experimental evidence and a model. In *Representing Musical Structure*, eds. P. Howell, R. West and I. Cross, Academic Press, London, pp 59 - 127.

Lerdahl, F., & Jackendoff, R. (1983). *A Generative Theory of Tonal Music.* Cambridge, MA: MIT press.

Longuet-Higgins, H. C. (1976). The perception of melodies. *Nature*, 263, 646 - 653.

Longuet-Higgins, H. C., & Lee, C. S. (1982). The perception of musical rhythms. *Perception*, 11, 115 - 128.

Longuet-Higgins, H. C., & Lee, C. S. (1984). The rhythmic interpretation of monophonic music. *Music Perception*, 1, 424 - 441.

MacDougall, R. The structure of simple rhythm forms. *Psychological Review, Monograph*, 4, 309 - 416.

Martin, J. G. (1972). Rhythmic (hierarchical) versus serial structure in speech and other behavior. *Psychological Review*, 79, 487 - 509.

McAdams, S., & Bigand, E. (1993). Introduction to auditory cognition. In *Thinking in Sound: The Cognitive Psychology of Human Audition*, eds., S. McAdams and E. Bigand, Oxford University Press, pp 1- 9.

Meyer, L. B. (1956). *Emotion and Meaning in Music*. University of Chicago Press, Chicago.

Michon, J. A. (1985). The Compleat Time Experiencer. *In Time, Mind, and Behavior*, eds. J. A. Michon and J. L. Jackson. Springer-Verlag, pp 20 - 52.

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*, 63, 81 - 97.

Miller, G. A. (1994). The magical number seven, plus or minus two: some limits on our capacity for processing information. Partial reprint or original work published in *Psychological Review*, 63, 81 - 97, *Perception & Psychophysics*, 101, 343 - 352.

Mishima, J. (1965). Introduction to the morphology of human behavior. The experimental study of mental tempo. Tokyo publishing, Tokyo. As cited in *The Psychology of Music*, ed. D. Deutsch, Academic Press, London, page 153.

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

Narmour, E. (1992). The influence of embodied registral motion on the perception of higher-level melodic implication. In *Cognitive Bases of Musical Communication*, eds., M. R. Jones and S. Holleran, American Psychological Association, Washington, DC, pp 69 - 90.

Ockelford, A. (1991). The role of repetition in perceived musical structures. In *Representing Musical Structure*, eds. P. Howell, R. West, I. Cross, Academic Press, London, pp 129 - 160.

Osgood, C. E. (1952). The nature and measurement of meaning. *Psychological Bulletin*, 49, 197 - 237.

Osgood, C. E., & Suci, G. J. (1955). Factor analysis of meaning. *Journal of Experimental Psychology*, 50, 325 - 338.

Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The measurement of meaning.* University of Illinois Press, Urbana.

Palmer, C. (1996). On the assignment of structure in music performance. *Music Perception*, 14, 23 - 56.

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

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

Palmer, C., & Krumhansl, C. L. (1990). Mental representations for musical meter. *Journal of Experimental Psychology: Human Perception & Performance*, 16, 728 - 741.

Parkin, A. J. (1987). *Memory & Amnesia: An introduction.* Basil Blackwell Ltd, Oxford.

Parncutt, R. (1987). The perception of pulse in musical rhythm. In *Action and Perception in Rhythm and Music*, ed., A. Gabrielsson, Royal Swedish Academy of Music, pp 127 - 138.

Parncutt, R. (1994). A perceptual model of pulse salience and metrical accent in musical rhythms. *Music Perception*, 11, 409 - 464.

Parncutt, R. (1989). *Harmony: A Psychoacoustical Approach.* Springer-Verlag Berlin Heidelberg.

Peretz, I. (1993). Auditory agnosia: a functional analysis. In *Thinking in Sound: The Cognitive Psychology of Human Audition*, eds., S. McAdams & E. Bigand, Oxford University Press, pp 199 - 230.

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

Pick, A. D. (1979). Listening to melodies: perceiving events. In *Perception and its development: A tribute to Eleanor J Gibson*, ed., A. D. Pick, Lawrence Erlbaum Associates, Hillsdale, New Jersey, pp 145 - 165.

Posner, M. (1964). Rate of presentation and order of recall in immediate memory. *British Journal of Psychology*, 55, 303 - 306.

Povel, D. J. (1981). Internal representations of simple temporal patterns. *Journal of Experimental Psychology: Human Perception & Performance*, 7, 3 - 18.

Povel, D. J. (1984). A theoretical framework for rhythm perception. *Psychological Review*, 45, 315 - 337.

Povel, D. J., & Collard, R. (1982). Structural factors in patterned finger tapping. *Acta Psychologica,* 52, 107 - 123.

Povel, D. J., & Essens, P. (1985). Perception of temporal patterns. *Music Perception*, 4, 411 - 440.

Povel, D. J., & Okkerman, H. Accents in equitone sequences. *Perception & Psychophysics*, 30, 565 - 572.

Pratt, C. C. (1968). *The Meaning of Music: A Study in Psychological Aesthetics.* Johnson Reprint Corporation, New York, London.

Reber, A. S. (1973). Locating clicks in sentences: Left, center, and right. *Perception & Psychophysics*, 1, 133 - 138.

Reber, A. S., & Anderson, (1970). The perception of clicks in linguistic and nonlinguistic messages. *Perception & Psychophysics*, 1970, 8, 81 - 89.

Repp, B. (1992). Probing the cognitive representation of musical time: structural constraints on the perception of timing perturbations. *Cognition*, 44, 241 -281.

Restle, F. (1972). Serial patterns: the role of phrasing. *Journal of Experimental Psychology*, 92, 385 - 390.

Royer, F. L. (1967). Sequential complexity and motor response rates. *Journal of Experimental Psychology*, 74, 199 - 202.

Rosenthal, D. A. (1989). A model of the process of listening to simple rhythms. *Music Perception*, 6, 315 - 328.

Sadie, S. (1981). *The New Grove Dictionary of Music and Musicians*, London Macmillan.

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

Shepard, R. N. (1982). Structural representations of musical pitch. In *The Psychology of Music*, ed., D. Deutsch, Academic Press, San Diego, California, pp 343 - 390.

Shaffer, L. H., & Todd, N. P. (1987). The interpretative component in musical performance. In *Action and Perception in Rhythm and Music*, ed., A. Gabrielsson, Royal Swedish Academy of Music, pp 139 - 152.

Sloboda, J. A. (1977). Phrase units as determinants of visual processing in music reading. *British Journal of Psychology*, 1977, 68, 117 - 124.

Sloboda, J. A. (1985). *The Musical Mind: The Cognitive Psychology of Music.* Clarendon Press, Oxford.

Sloboda, J. A. (1988). Preface. In *Generative Processes in Music: The Psychology of Performance, Improvisation, and Composition*, ed., J. A. Sloboda, Oxford University Press, Oxford, pp v - xii.

Sloboda, J. A., & Gregory, A. H. (1980). The psychological reality of musical segments. *Canadian Journal of Psychology*, 34, 274 - 280.

Sloboda, J. A., & Parker, D. H. H. (1985). Immediate recall of melodies. In *Musical Structure and Cognition*, eds., P. Howell, I. Cross, R. West, Academic Press, London, pp 143 - 167.

Smith, J. D. (1997). The place of musical novices in music science. *Music Perception*, 14, 227 - 262.

Smith, C., & Cuddy, L. (1989). Effects of metric and harmonic rhythm in the detection of pitch alterations in musical sequences. *Journal of Experimental Psychology: Human Perception & Performance*, 15, 457 - 471.

Speer, J. R., & Meeks, P. U. (1985). School children's perception of pitch in music. *Psychomusicology*, 5, 49 - 56.

Steedman, M. J. (1977). The perception of musical rhythm and metre. *Perception*, 6, 555 - 569.

Sternberg, S., & Knoll, R. L. (1994). Perception, production, and imitation of time ratios by skilled musicians. In *Musical Perceptions*, eds., R. Aiello and J. A. Sloboda, Oxford University Press, pp 240 - 257.

Stoffer, T. H. (1985). Representation of phrase structure in the perception of music. *Music Perception*, 3, 191 - 220.

Stoll, G., & Parncutt, R. (1987). Harmonic relationship in similarity judgements of nonsimultaneous complex tones. Acustica, 63, 111 - 119.

Summers, J., Bell, R., & Burns, B. (1990). Perceptual and motor factors in the imitation of simple temporal patterns. *Psychological Review*, 50, 23 - 27.

Swain, D. (1986). The need for limits in hierarchical theories of music. *Music Perception*, 4, 121 - 148.

Tan, N., Aiello, R., & Bever, T. G. (1981) Harmonic structure as a determinant of melodic organization. *Memory & Cognition*, 9 533 - 539.

Temperley, N. M. (1963). Personal tempo and subjective accentuation. *Journal of General Psychology*, 68, 267 - 287.

Terhardt, E. (1987). Gestalt principles and music perception. In *Auditory Processing of Complex Sounds*, eds., W. A. Yost, C. S. Watson, Erlbaum, Hillsdale, New Jersey, pp 157 - 166.

Thompson, W. F. (1993). Pitch pattern, durational pattern, and timbre: a study of the perceptual integration of auditory qualities. *Psychomusicology*, 12, 3 - 21.

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

Thorpe, L. A., Trehub, S. E., (1989). Duration illusion and auditory grouping in infancy. *Developmental Psychology*, 25, 122 - 127.

Thorpe, L. A., Trehub, S. E., Morrongiello, B. A., & Bull, D. (1988). Perceptual grouping by infants and preschool children. *Developmental Psychology*, 24, 484 - 491.

Tighe. T. J., & Dowling, W. J. (1993). Preface. In *Psychology and Music: The Understanding of Melody and Rhythm*, eds, T. J. Tighe and W. J. Dowling, Lawrence Erlbaum Associates, Hillsdale, New Jersey, Hove and London, pp vii - viii.

Todd, N. M. (1985). A model of expressive timing in tonal music. *Music Perception*, 3, 33 - 58.

Todd, N. M. (1994). The auditory "Primal Sketch": a multiscale model of rhythmic grouping. *Journal of New Music Research*, 23, 25 - 70.

Trehub, S. E., & Trainor, L. J. (1993). Listening strategies in infancy: the roots of music and language development. In *Thinking in Sound: The Cognitive Psychology of Human Audition*, eds., S. McAdams & E. Bigand, Oxford University Press, pp 278 - 327.

Vitz, P. C., & Todd, T. C. (1969). A coded element model of the perceptual processing of sequential stimuli, *Psychological Review*, 76, 433 - 449.

Vos, P. G., Dijk, A. V., & Schomaker, L. (1994). Melodic cues for metre. *Perception*, 1994, 23, 965 - 976.

Vos, P. G., & Troost, J. M. (1989). Ascending and descending melodic intervals: statistical findings and their perceptual relevance. *Music Perception*, 6, 383 - 396.

Waterworth, J. A. (1983). Effect of intonation form and pause duration of automatic telephone number announcements on subjective preference and memory performance. *Applied Ergonomics*, 14, 39 - 42.

Watt, R. J. (1995). Perception of meaning in Wagner's Siegfried by naïve subjects. Paper presented at the *DGM/ESCOM annual symposium*, ed. G. Kleinen, University of Bremen.

Watt, R. J., & Ash, R. L. (1996). An experimental investigation of the psychological content of music. In *Fourth International Conference on Music Perception and Cognition*, eds., B. Pennycook, E. Costa-Giomi, McGill University, Montreal, pp 161 - 166.

Wedin, L. (1972). A multidimensional study of the perceptual-emotional qualities in music. *Scandanavian Journal of Psychology*, 13, 241 - 257.

West, R., Howell, P., & Cross, I. (1991). Musical Structure and Knowledge Representation. In *Representing Musical Structure*, eds. P. Howell, R. West, I. Cross, Academic Press, London, pp 1 - 30.

Wertheimer, (1927). Untersuchungen zur Lehre von der Gestalt, II. Psychologische Forschung, 4, 301 - 150. As cited in *Gleitman*, H. Gleitman, Psychology, 2^{nd} edition, Norton & Company, New York, London, 1986, pp 183 - 186.

Wing, A. M., & Kristofferson, A. B. (1973a). The timing of inter-response intervals. *Perception & Psychophysics*, 13, 455 - 460.

Wing, A. M., & Kristofferson, A. B. (1973b). Response delays and the timing of discrete motor responses. *Perception & Psychophysics*, 14, 5 - 12.