Contextual Effects in Pitch Processing: Investigating Neural Correlates Using Complementary Methodologies

Catherine M. Warrier
Department of Psychology
McGill University

Montreal, Quebec
Canada

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements of the degree of PhD.

Submitted May 2000

0-612-69945-5

Dedication:

This work is dedicated to my husband Kishore for his constant and boundless love and support, and to my wonderful parents for their encouragement of my interests in both science and music.

Contextual Effects in Pitch Processing:
Investigating Neural Correlates Using Complementary Methodologies

Table of Contents:

Table of Contents continued

Chapter 5: Investigating the Separation of Lyrics and Melody in Song Through PET

Chapter 6: General Discussion

Abstract

This thesis includes four studies investigating neural correlates underlying pitch perception, and effects of tonal context on this percept. Each study addressed the issue from a unique methodological perspective. The first study confirmed that tonal context can affect the way a tone's pitch is perceived. In this study, normal listeners made pitch discriminations between tones varying in pitch and/or timbre, a difficult task when presented in isolation. Increasing tonal context increased performance, with melodic context providing the most facilitation.

A similar task was presented to patients with unilateral focal excisions in the temporal lobe. Patients with right but not left temporal lobe lesions were impaired at using melodic cues to facilitate performance. Posterior extent of the lesions did not affect results, implying that right anterior temporal regions can process pitch information relative to tones heard previously. A functional magnetic resonance imaging (fMRI) study using a similar task with normal listeners found converging evidence. Melodic context produced the most activity in right anterior superior temporal gyrus (STG), as well as the most facilitation behaviorally.

A positron emission tomography study investigating neural processing of song stimuli broadened the investigation to include a comparison between musical and linguistic processing. Left frontal and temporal structures known to be involved in language processing were active when subjects attended to song lyrics, and right temporal-lobe structures were again implicated in melodic processing, suggesting that a song's lyrics and melodies are processed separately.

These studies find pitch processing in tonal contexts to involve right temporal-lobe structures. The right anterior STG in particular appears to be involved in processing pitch relative to previously heard tones. This suggests that the right anterior STG processes tones with respect to their tonal context, which entails holding contextual tones in memory while processing subsequent tones. This region has connections to right dorsolateral frontal areas previously implicated in tonal working memory, possibly providing a mechanism for holding contextual tones in memory. Supporting this theory, all contextual conditions in the fMRI study produced activity in right dorsolateral frontal cortex.

Résumé

Cette thèse inclut quatre études portant sur les corrélats neuronaux de la perception de la hauteur et sur le contexte tonal de sa perception. Chaque étude a abordé cette question avec une méthodologie unique. La première étude a confirmé que le contenu tonal peut affecter la manière dont la hauteur d'un son est perçue. Dans cette étude, des sujets normaux ont fait une tâche de discrimination de la hauteur, pour des tons variant en hauteur et ou en timbre, une tâche difficile considérant qu'elle était presentée isolément. L'augmentation du contexte tonal a amélioré la performance. Le contexte mélodique a le plus facilité cette amélioration de la performance.

Une deuxième tâche a été présentée à des patients ayant une excision focale et unilatérale dans le lobe temporal du cerveau. La performance pour les patients ayant les lésions du côté droit et non du côté gauche fut dimimuée lorsqu'ils devaient utiliser des indications mélodiques. L'étendue postérieure des lésions n'a pas affecté les résultats, impliquant que les régions antérieures temporales peuvent traiter la hauteur relative aux tons endendus auparavant.

Une étude d'imagerie par résonance magnétique fonctionnelle (IRMF) utilisant une tâche similaire avec des sujets normaux a permis de trouver des résultats convergente. Le contexte mélodique a généré le plus d'activité dans la région antérieure du gyrus temporal supérieur (GTS) tout en facilitant le plus les tâches comportementales.

Une étude de tomographie par émission de positons portant sur les processus neuronaux impliqués dans la perception des chansons a approfondi l'étude en incluant une comparaison entre les processus musicaux et linguistiques. Les structures temporales et frontales du côté gauche, probablement impliquées dans le traitement du langage, furent activées lorsque les sujets prêtaient attention aux paroles des chansons. Lorsque les sujets prêtaient attention à la mélodie, les structures du lobe temporal droit étaient activées, suggérant que les paroles de chansons et les mélodies sont interpretées et traitées séparément par différentes régions du cerveau.

Ces études démontrent que les structures du lobe temporal droit sont impliquées dans le traitement de la hauteur dans un contexte tonal. Le partie antérieure du GTS du côté droit, en particulier, semble être importante pour percevoir la hauteur relative à des tons entendus au préalable. Ceci suggère que cette région du cerveau traite des sons par rapport à leur contexte tonal, ce qui implique de tenir en mémoire des tons dans leur contexte tout en prêtant attention aux tons subséquents. Le partie antérieure du GTS du côté droit a des connections aux régions frontales et dorsolatérales du côté droit déjà impliquées dans la mémoire de travail tonal, possiblement un méchanisme pour retenir des tons dans leur contexte en mémoire. Cette théorie est appuyée par les résultats de l'étude d'IRMF où toutes les conditions contextuelles ont généré de l'activité dans le cortex frontal dorsolatéral du côté droit.

Contributions of Authors on published / submitted articles

McGill University requires that in the case of papers included in a thesis that have been published or submitted for publication, as have chapters 2-5 of this thesis, an explicit statement of the contributions of each author must be made.

Chapter 2: Influence of Tonal Context and Timbral Variation on Perception of Pitch

Catherine M Warrier and Robert J Zatorre

Catherine M Warrier:

- Outlined rationale / procedure of study in conjunction with R Zatorre
- Created and tested all stimuli
- Wrote stimulus presentation / data collection program (with P Ahad's help)
- Recruited, scheduled, and tested subjects
- Analyzed data
- Interpreted results in conjunction with R Zatorre
- Presented data as conference talk and wrote paper for journal submission

Robert J Zatorre:

- Participated in outlining rationale / procedure of study
- Participated in interpretation of results
- Gave valuable input throughout experimental process
- Proofed drafts of journal submission

Chapter 3: The Role of Right Temporal Cortex in the Effect of Melodic Context

Catherine M Warrier and Robert J Zatorre

Catherine M Warrier:

- Outlined rationale / procedure of study in conjunction with R Zatorre
- Created and tested all stimuli
- Wrote stimulus presentation / data collection program (with P Ahad's help)
- Scheduled and tested subjects
- Analyzed data
- Interpreted results in conjunction with R Zatorre
- Presented data as conference talk and wrote paper for journal submission

Robert J Zatorre:

- Participated in outlining rationale / procedure of study
- Participated in interpretation of results
- Gave valuable input throughout experimental process
- Proofed drafts of journal submission

Chapter 4: Right Anterior Superior Temporal Gyrus Implicated in Tonal Context Effects

Catherine M Warrier, Pascal Belin, Isabelle Merlet, and Robert J Zatorre

Catherine M Warrier:

- Generated rationale, general procedure of study
- Obtained approval from MRI working committee / MNI ethics committee
- Created and tested all stimuli
- Wrote stimulus presentation / data collection program (with P Ahad's help)
- Recruited and scheduled subjects
- Ran testing of subjects; stimulus presentation and behavioral data collection
- Analyzed functional and behavioral data
- Interpreted results in conjunction with R Zatorre
- Presented results as conference poster and wrote paper for journal submission

Pascal Belin:

- Participated in discussions of procedure
- Volunteered as pilot subject
- Participated heavily in functional data analysis of pilot subjects
- Guided me through the complex logistics of running an fMRI study
- Aided in post-hoc functional data analyses

Isabelle Merlet:

- Participated in discussions of procedure
- Helped refine stimulus presentation paradigm during extensive pilot testing

Robert J Zatorre:

- Participated in discussions of rationale / procedure
- Provided funding for use of fMRI facilities
- Participated in interpretation of results
- Proofed drafts of journal submission

Chapter 5: Investigating the Separation of Lyrics and Melody in Song through PET

Catherine M Warrier, Robert J Zatorre, Isabelle Peretz, and Mireille Besson

Catherine M Warrier:
- Outlined rationale / procedure of study in conjunction with other authors
- Obtained approval from PET working committee and MNI ethics committee
- Worked stimuli into proper format for use in this study
- Scheduled scanner time
- Ran testing of subjects; stimulus presentation and recording of behavioral data
- Analyzed data
- Interpreted results in conjunction with R Zatorre
- Presented data as conference poster and wrote paper for journal submission

Robert J Zatorre:
- Participated in discussions of rationale / procedure for study
- Helped recruit and schedule subjects
- Provided funding for use of PET scanner
- Read task instructions during scanning procedures
- Participated in interpretation of results
- Proofed drafts of journal submission

Isabelle Peretz:
- Participated in discussions of rationale / procedure for study
- Provided use of lab to work stimuli into usable form
- Helped recruit and schedule subjects
- Read task instructions during scanning procedures

Mireille Besson:
- Participated in original discussion of rationale / procedure for study
- Provided opera stimuli

Acknowledgements:

I would first like to thank my supervisor, Robert Zatorre, for always being there when I needed guidance, as well as giving me the freedom to work out problems and discover things on my own. Thank you also to Pierre Ahad for his expert technical assistance in guiding me through the complexities of creating stimulus presentation programs in MAPLE, and to Giovanna Locascio for handling my many beurocratic difficulties with ease. I would like to thank the MNI Neuropsychology department in general, and my labmates in particular, for the creative and supportive working environment we all enjoy. A special thank you goes to Patrick Bermudez for his companionship, support, and unending generosity in all things. To Anne, Laila, Jim, Stéphane and Patrick, thank you for your friendship and inspiration, making my time in Montreal easier and enjoyable in so many ways. Last but certainly not least, I would like to thank my husband Kishore for just about everything, but particularly for his patience throughout this past year.

List of Abbreviations:

☞	Cent: Logarithmic difference in frequency
AA2, AA3	Auditory association areas located anteriorly on superior temporal gyrus
ANOVA	Analysis of Variance
BA##	Brodmann's areas
BOLD	Blood Oxygenation Level Dependent
CBF	Cerebral Blood Flow
DAT	Digital Audio Tape
dB SPL	Sound Pressure Level in decibels
EEG	Electroencephalography
ERP	Evoked Response Potential, an EEG technique
F0	Fundamental Frequency
fMRI	Functional Magnetic Resonance Imaging
FWHM	Full Width Half Maximum
Hz	Hertz, measured in cycles per second
IQ	Intelligence Quotient
ISI	Interstimulus Interval
LIF	Left Inferior Frontal cortex
LPC	Late Positive Component, an auditory evoked potential
LT	Group of patients with left temporal lobe excisions
Mel_CC	Test condition in Chapter 4: Monitor melodies while both semantics and melody are congruous
Mel_IC	Test condition in Chapter 4: Monitor melodies while melodies only are incongruous
Mel_II	Test condition in Chapter 4: Monitor melodies while both semantics and melody are incongruous
MRI	Magnetic Resonance Imaging
N400	Auditory evoked potential seen as a negative peak of around 400 msec latency
NC	Control group of neurologically normal participants
PET	Positron Emission Tomography
RT	Chapter 3: Group of patients with right temporal lobe excisions Chapter 4: Reaction time measured in milliseconds
Sem_CC	Test condition in Chapter 4: Monitor semantics while both semantics and melody are congruous
Sem_IC	Test condition in Chapter 4: Monitor semantics while semantics only are incongruous
Sem_II	Test condition in Chapter 4: Monitor semantics while both semantics and melody are incongruous
STG	Superior Temporal Gyrus
WAIS-R	Wechsler Adult Intelligence Scale - Revised

Chapter 1

General Introduction

General Overview

A vital role of our perceptual system is to be able to interpret the same stimulus differently under differing circumstances. For example, the same sound can mean something entirely different when heard in a different context. This is intuitively true for verbal, environmental, and ambient sounds, but also applies to tonal sounds, in that surrounding tones may influence the perception of a tone's pitch (Deutsch, 1972a,b, 1982; Deutsch & Roll, 1974; Dewar et al., 1977; Krumhansl, 1979; Krumhansl & Castellano, 1983).

In this thesis, the neural correlates underlying pitch perception and the effect of tonal context on this percept will be addressed in a series of four research studies. Each of these studies utilizes a different neuropsychological method, adding a unique perspective on various facets of the research question. This introduction will begin by addressing the issue of how tonal and musical stimuli are useful when investigating cognitive questions relating to organization of the auditory perceptual system, then describe some basic concepts about tonal stimuli. An introduction to current literature on contextual effects of pitch processing and a review of the relevant neuropsychological literature follows. An account of the advantages and disadvantages of the different methodologies used in this thesis are then addressed, in addition to how each method complements the others. This chapter ends with brief descriptions of each of the four studies reported, and how they relate to and progress from each other.

Benefits of using tonal stimuli

Tonal and musical stimuli enable the study of some interesting aspects of the organization of the auditory perceptual system that are difficult to test with other auditory stimuli. For example, a single tone is generally devoid of the extraneous meanings and associations that are bound to words and environmental sounds. This feature of tones allows a more direct investigation of purely auditory phenomena by avoiding some confounds associated with other types of auditory stimuli. By carefully manipulating both stimulus and task, it is possible to use tonal and musical stimuli to study many cognitive processes such as perception and memory.

By combining tones in different ways, varying degrees of tonal context may be generated, an important feature of the studies reported here. For example, tones can be presented in isolation without any surrounding tones, in alternating or repeating sequences, or in random or melodic sequences. Melodies in particular are a special form of tonal stimuli that allow unique insights into auditory processes. The rule-based system of tonality can guide melodic manipulations to exploit different aspects of tonal perception for study (e.g. Krumhansl, 1990).

While the study of tonal and musical stimuli is crucial for a complete understanding of the auditory system, it should not be pursued to the exclusion of other aspects of auditory perception, such as verbal language. By comparing and contrasting information gained through research with different types of auditory stimuli, a more comprehensive view of this system can be attained. It is for this reason that the study reported in Chapter 5, which combines linguistic and melodic features in song stimuli, is included in this thesis. By carefully manipulating the linguistic and melodic aspects of the stimuli separately and in unison, then combining the task with a functional neuroimaging technique, it was possible to directly compare the neural substrates involved in musical and linguistic processing.

Aspects of tonal stimuli

Quantifying pitch

Changing the physical attributes of a tone affects its perceptual quality. Changes in frequency lead perceptually to differences in pitch, while changing the spectral attributes of a tone affects its timbre, but there is not always a one-to-one correspondence between the physical and perceptual attributes. For example, at lower frequencies, increasing the frequency by 100 Hz changes the pitch of a tone considerably, while the same amount at higher frequencies results in relatively minimal pitch change. A psychophysical scale of pitch has been described, using the term “mel” for units of subjective pitch, in which doubling the number of mels doubles the pitch, and a 1000 Hz tone is defined as having a pitch of 1000 mels (Stevens et al., 1937; summarized in Rossing, 1990; Handel, 1993). This scale was created by having listeners tune a comparison tone to half the pitch of a standard tone. Inspection of the relationship between frequency and mels shows that subjective pitch does not increase as quickly as frequency. For example, a nine-fold increase in frequency (1000 vs. 9000 Hz) is perceived as only a three-fold increase in pitch (1000 vs. 3000 mels).

Another way of measuring pitch is related to the musical scale. In the Western system of tonality, an interval of a musical octave is made by doubling the frequency of a tone. The octave is then divided into twelve semitones, with each semitone considered an equal step in pitch. In the equal tempered scale, the octave is subdivided into these twelve semitones by logarithmically equal steps in frequency, with the frequency ratio between the two tones of a semitone interval equal to $2^{1/12}$, or 1.05946. These perceptually equal steps in pitch can be measured by another psychophysical measure called cents. Each cent is defined as 1/100th of a musical semitone, with 100 cents to a semitone, and 1200 cents to an octave (summarized in Rossing, 1990). The equation for converting a frequency ratio to an interval in cents is:

$$\text{Cent\_Interval} = (1200 / \log 2) \log \text{Frequency\_ratio}$$

As an example, the interval of a musical fifth, which is 7 semitones apart, has a frequency ratio of 1.498 in the equal tempered scale. Putting this frequency ratio into the equation produces $(1200 / \log 2) \log 1.498 = 700$ cents. Frequency differences producing intervals less than a semitone are also easily quantified into cents using this equation. Since any frequency interval is easily converted into cents, a measure meaningful for quantifying perceptual differences in pitch relative to known musical intervals, this thesis uses cents to describe differences in frequency.

Attributes of complex tones

Most of the stimuli described in this thesis consist of complex periodic tones. Complex tones contain components at multiple frequencies, with each frequency component, or harmonic, being an integer multiple of the lowest, or fundamental frequency (F0). This feature of complex tones is common to many musical instrument sounds. The pitch of complex tones usually corresponds to the pitch of the F0. However, if this fundamental frequency component is filtered out, such that all harmonics are integer multiples of an absent frequency, the pitch of that tone still corresponds to the missing F0 of that tone (summarized in Moore, 1989). This perceptual phenomenon is termed virtual, or residue pitch. However, although the pitch of these tones remains the same, removing the fundamental, or any harmonics of a tone, may affect the timbre, or sound quality, of that tone (summarized in Moore, 1989).

The timbre of a tone is described as being "that attribute of auditory sensation in terms of which a listener can judge that two sounds similarly presented and having the same loudness and pitch are dissimilar" (American National Standards Institute, 1960). For example, a piano and an organ can both play the same note at the same intensity level, but because they have different timbral properties, it is still possible to distinguish the piano tone from the organ tone. Many different physical properties can alter the timbre of a sound. These properties can be grouped into two categories: 1) spectral composition, comprised of the number of harmonics, these harmonics' relative intensities, as well as the pattern of spectral change across the duration of the tone, and 2) amplitude envelope, consisting of the rate of onset of a tone, the duration of its steady state, and its decay rate.

Multidimensional scaling studies have been useful in determining the perceptual significance of different timbral properties. Most of these studies conclude that the two most salient timbral properties correspond to a spectral and temporal dimension (Miller, 1975; Grey, 1977; Ehresman & Wessel, 1978; Wessel, 1979; Krumhansl, 1989; McAdams et al., 1995; Samson et al., 1997). In these studies, the spectral dimension is associated with the "brightness" of timbre, which, like pitch, relates directly to the frequency components of the harmonics. Increasing the number of harmonics in a tone, or varying their intensities so that relatively more energy is found in the higher harmonics, can increase that tone's "brightness." Two of these studies also found spectral variation across the duration of the tone as a third significant dimension (Grey, 1977; Krumhansl, 1989). Due to the multidimensionality of timbre, the intricacies of its analysis and synthesis can be very complex (See Risset & Wessel, 1999 for a review). These intricacies will not be reviewed here, since the timbral manipulations included in this thesis, simple manipulations of the relative intensities of a fixed number of harmonics, were straightforward and secondary to an insight into the perceptual properties of pitch.

Pitch chroma

In addition to the increasing pitch "height" perceived with increasing frequency, there is a cyclical aspect to pitch, called "chroma" (See Krumhansl, 1990; Deutsch, 1999 for reviews). The idea of chroma can be seen in the repetitive nature of musical scales; the same note names (A through G#) are repeated with each doubling of frequency, resulting in a complete set of note names per octave. Therefore, notes whose F0s are integer multiples of each other share the same note names. Musically speaking, these notes are separated by intervals of one or more octaves. Two tones that create an octave interval are not only musically related, but due to their whole number frequency ratios (2:1, 3:1, 4:1, etc.) they are very similar physically as well. When the F0 of a tone is an integer multiple of another, the harmonics of the higher tone will map directly onto the harmonics of the lower tone (See Rossing, 1990 for review).

Take for instance A3, a note in the third octave with chroma name "A" and an F0 of 220. Its first 4 harmonics are 220, 440, 660, and 880 Hz. A4, which is one octave higher, has an F0 of 440. Its first two harmonics, 440 and 880 Hz, are exactly the same as the 2nd and 4th harmonics of A3. As frequency ratios get more complex, fewer harmonics overlap between the two notes, increasing the dissonance between them (summarized in Krumhansl, 1990; Burns, 1999). In addition, the complexity of frequency ratios between two notes also affects their relative importance within a particular musical key. Tones with simpler ratios to the tone the key is based on (called the tonic in the Western system of tonality) enjoy a more important role within that key. A basic description of this tonality system follows.

Perception of pitch within the Western system of tonality

The Western system of tonality is the musical system most commonly encountered in North American culture, and was used in creating tonal contexts for the purposes of this thesis. This system can be described as a complex grammar for arranging notes both sequentially (melodically) and simultaneously (harmonically) to form musical progressions. Its grammar derives from a combination of hierarchical systems guiding melodic, harmonic, and key progression (Summarized in Krumhansl, 1990; Bigand, 1993; Handel, 1993). A schematic diagram of these hierarchies is included in Figure 1.

As indicated on this diagram, the building blocks of the Western system of tonality are the twelve notes of the chromatic scale. Individual major and minor scales are created with subsets of seven of these twelve notes, with 24 possible scales. Within each scale, or key, some notes are more important than others. The note the scale starts on and is named after (e.g. "C" in the C major scale), also called the tonic, is always the most important note. In practice, this note is emphasized through a variety of different ways. Examples of how it is emphasized include frequent occurrences, dynamic and/or rhythmic accentuation, and prominent placement within the phrases of a piece. An important musical aspect that comes out of this system is that tones of less relative importance within a key tend to lead to tones of higher importance. This idea has been described in terms of stability and tension; less important tones sound more unstable, creating musical tension which is then resolved by tones of higher stability (e.g. Bigand, 1993; Handel, 1993).

Another source of rules in this system governs the relatedness of different keys. The circle of fifths depicted in Figure 1 is an effective way of illustrating this system, where keys that are close together on the circle are closely related, and keys that are distant from each other on the circle are distantly related. This image is called the circle of fifths because each key is an interval of a fifth away from the two adjacent keys on the circle. The relatedness of different keys comes into play when musical compositions present sections of a piece in different keys, a feature common to many standard musical forms, such as the sonata form.

Many perceptual studies have shown that these hierarchies are perceived by musicians and non-musicians alike, and that the system is learned in childhood through either explicit or implicit exposure (e.g. Speer & Adams, 1985 as described in Krumhansl, 1990, Chapter 2; see Dowling, 1999 for a review). For example, one method of investigating internal representations of these tonal hierarchies involves establishing a

Schematic Diagram of the Western System of Tonality

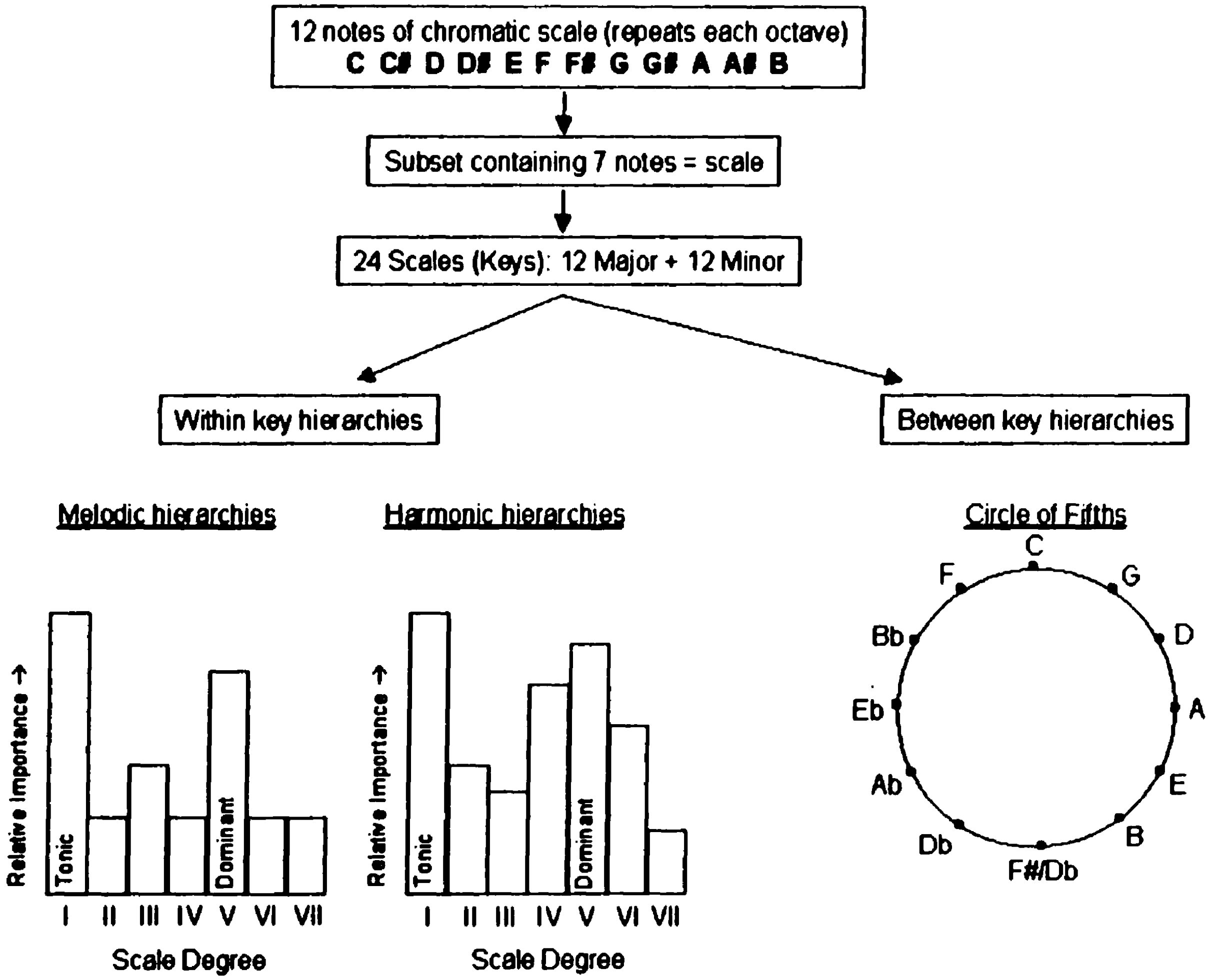

Figure 1: Schematic diagram of the western tonal system, adapted from Bigand (1993). Each scale, or key, is made up of 7 out of 12 possible notes, which repeat every octave. Three types of hierarchies determining the relative importance of musical events are illustrated. For ease of illustration, only the hierarchies governing major scales are shown. The melodic hierarchy indicates the relative importance of each note within a melody. Scale degrees are shown by roman numerals starting with the tonic, such that in the key of C Major, I=C, II=D, III=E, etc. The harmonic hierarchy indicates the relative importance of chords within a harmonic progression. Scale degrees of harmonic hierarchies each indicate the chord triad starting on the note indicated by a roman numeral, such that in the key of C Major, I=C+E+G, II=D+F+A, etc. The circle of fifths represents the relative closeness of keys to each other, with closely related keys located adjacent to each other on the circle, and remotely related keys located further apart.

key by presenting a scale (Krumhansl & Shepard, 1979; Krumhansl & Kessler, 1982), or tonic triads and chord cadences (Krumhansl & Kessler, 1982), and then a "probe tone" for which the listeners must rate how well that tone either completes the scale or fits with the tonal context. In these studies, listeners' ratings tend to increase with increasing relative importance of the probe tone as described by the rules of the tonality system. For example, the tonic (I) is always given the highest ratings. Other methods such as rating phrase endings (Boltz, 1989), and measuring reaction time when judging key membership (Janata & Reisberg, 1988) also show evidence that listeners have implicit knowledge of the hierarchies inherent to the tonal system, and perceive musical stimuli relative to this knowledge (See Bigand, 1993 for review).

Musical expectation

Two types of musical expectations have been described, veridical and schematic, that predict musical events such as chords, tones, or melodic lines, based on different types of information (Bharucha, 1994). Veridical expectancy predicts musical events based on memories of previous presentations of that particular piece of music, thus predicting events that actually follow. Schematic expectancy, on the other hand, predicts "culturally generic" musical events that typically follow, based on a mental representation, or schema, built up from exposure to many examples of musical events based on a particular system of tonality. Schematic and veridical expectations do not have to coincide; it is possible that a schematically expected tone will not actually occur, thus conflicting with veridical expectancy.

Musical schemas can be created either explicitly through musical training, or implicitly through listening to the music that is so pervasive in society (from radio, movies, TV, grocery stores, elevators, shopping malls, etc), creating an intuitive understanding of the rules governing the tonal system. Within this system, certain combinations of melodies, harmonies, and rhythms set up expectations in our mind as to what could or should follow (Bharucha & Stoekig, 1986; Bharucha, 1994). These schemas will be culturally specific, in that the music of one culture may follow different tonal rules than another, such as the Western tonal system compared to the systems governing Indonesian gamelan or Indian karnatic or hindustani music.

Psychological evidence has indicated that schematical expectancies of Western listeners relates to musical events predicted by the rules of the Western system of tonality (Cuddy et al., 1981; Bharucha & Stoekig, 1986; Schmuckler, 1989; Schmuckler, 1997). For example, one study had listeners rate the strength and specificity of expectations formed by listening to musical excerpts, then rate how well different continuations of these excerpts conformed to their expectations (Schmuckler, 1989). Melodic and harmonic aspects of these continuations were manipulated both separately and together in the first three experiments. Continuations played in accordance with the melodic and harmonic hierarchies outlined by music theorists (summarized in Figure 1) were rated as having higher conformance to expectation than those that violated these hierarchies. Melodic and harmonic expectations were found to be independent and additive. A fourth experiment reported in this study had pianists play out their expectancies to the same excerpts on a piano. These continuations were highly correlated with the expectancy ratings given in experiments 1-3, thus providing converging evidence.

Other studies find faster and/or more accurate responses to musical events when they are predicted by rules of tonality than when they violate these rules, implying that the listener forms an expectation conforming to these rules which then prepares the listener for the expected event (Cuddy et al., 1981; Bharucha & Stoekig, 1986). For example, a study by Bharucha and Stoekig (1986) asked listeners to judge whether the second chord in a two-chord sequence was presented in or out of tune (target chords were mistuned on half of the trials). In half of the sequences, the chords followed a common harmonic progression as defined by the tonal system, thereby fulfilling expectations, whereas the remaining sequences were unexpected within the tonal system. Musician and non-musician listeners alike were both more accurate and faster to respond when the chords were expected within the tonal system.

Therefore, an important aspect of tonal systems is that they structure tones into progressions that create expectations within the listener's mind. These expectations conform with the rules of the tonal system, and can then either be upheld or violated by the actual musical event. Expectation appears to serve a preparatory role in perceiving subsequent events.

Contextual effects on pitch processing

The main cognitive process investigated in this thesis is the effect of tonal context on pitch perception. Previous studies addressing this topic have established that the pitch of a tone can be affected by its surrounding tones (Deutsch, 1972a,b; Deutsch, 1982; Deutsch & Roll, 1974; Dewar et al., 1977; Krumhansl, 1979; Krumhansl & Castellano, 1983; Rakowski, 1990; Umemoto, 1990). In addition to the expectation created by certain contexts, as described above, it appears that contextual tones create a tonal reference point that may influence how the pitch of a tone is perceived. Most musical systems, including the Western tonal system, are based on the relationship of frequency ratios between tones. A two-tone interval has a specific ratio between F0's that can be either in or out of tune within that system. Hearing more notes creates more of a tonal reference point by which one can judge whether each tone is in or out of tune with the rest of the notes present (e.g. Rakowski, 1990).

For example, one study asked listeners to detect a quarter-tone mistuning in one of a sequence of six tones (Umemoto, 1990). The tone sequences were either arranged according to the western system of tonality, or not (atonal). Both musicians and non-musicians were better able to detect mistunings when the system of tonality was observed. Since the mistunings were identical in each context, this result suggests that the pitches of these tones were perceived differently in the different contexts.

Another study investigated differences in memory for pitches presented in different contexts (Krumhansl, 1979). In this study, listeners were asked to compare the pitch of the first and last tones of eight- and ten-tone sequences. In half of the sequences, the intervening tones evoked the key of C major, while the other sequences evoked no key at all (atonal context). Subjects rated the pitch of the test tones on a scale from "very sure same" to "very sure different." On same-pitch trials presented in the C major context, important tones of the C major tonal hierarchy (C, E, and G) were judged "same" most confidently, and tones not belonging to the C major scale (e.g. F#) were judged with the least confidence. This difference in confidence ratings was significantly less pronounced in the atonal context, again indicating that the pitch of the same tones can be perceived differently in different contexts.

Related auditory perceptual phenomena will be addressed in conjunction with the concept of contextual effects in detail in later chapters. For example, Chapters 2 and 3 deal with contextual effects on perceptual constancy, the ability to determine that a stimulus stays the same under differing conditions. In addition, comparisons between tonal and verbal processing of sung stimuli will be discussed in Chapter 5. A discussion of the neuropsychological literature on perceptual processing of tonal and melodic stimuli follows.

Review of neuropsychological studies on tonal and melodic processing

Neuropsychology, or cognitive neuroscience, a field linking behavior, cognition and perception to brain function, has studied the relationship between music and the brain for many years. This research has found no particular cortical region involved in global music processing per se. Instead, studies have revealed that musical processing recruits different neural systems distributed throughout the brain depending on the specific aspect of musical processing investigated. For example, melody recognition, timbre discrimination, and rhythm perception, all distinct aspects of music processing, each appear to involve different neural systems (Samson & Zatorre, 1994; Lechevalier et al., 1995; Penhune et al., 1999). It is therefore very important when discussing the neuropsychological literature on music perception to be specific as to the exact musical process under study. As this thesis focuses on the perception of pitch and pitch sequences, the literature relevant to pitch processing will be reviewed here.

Amusia

Historically, there have been some reports of patients with brain damage causing a dissociation of musical function from other processes, particularly language. This type of dissociation is an important finding in neuropsychological research because it implies that the neural structures involved in music processing are separate from those involved in linguistic processing. A famous example of this sort of dissociation of function involved a composer who sustained damage to his left temporal and temporoparietal regions through a stroke (Luria et al., 1965). Although this damage caused severe impairments in the man's language perception and production, his musical abilities remained intact. In fact, the music he composed after his stroke was considered some of his best by fellow musicians. The opposite pattern of function in which a patient loses his/her musical abilities while linguistic abilities are spared has been documented in other patients (e.g. Peretz et al., 1994). The neurological disorder involving specific impairments in musical ability while sparing other functions is termed amusia, and the study of amusic patients plays an important role in dissociating musical function from other cognitive and perceptual processes.

Some recent cases of amusic patients have been carefully and systematically studied by Isabelle Peretz and her colleagues, uncovering important information about the organization of neural systems underlying music perception (Peretz et al., 1994; Peretz, 1996; Peretz et al., 1997). These cases provide evidence of dissociations not only between the processing of musical sounds from linguistic or environmental sounds, but of dissociations between sub-components of music processing as well. For example, a dissociation between melodic and rhythmic processing was observed in the two patients CN and GL (Peretz et al., 1994).

Due to the rareness of this disorder, and the fact that the neural damage incurred by amusic patients is usually quite dispersed, it is generally difficult to use case studies of these patients to localize musical perceptual systems within the brain. However, the testing of patients with focal cortical lesions does allow the localization of cortical structures involved in these systems. By testing patients with known focal lesions on various musical tasks, it is possible to determine if the damaged area is important in performing those tasks or not (see below for more detail on this technique).

Lesion studies

The first study investigating music perception in this way tested patients with unilateral anterior temporal lobectomies on a series of musical tasks taken from the

Seashore Measures of Musical Talents: discrimination of pitch, loudness, rhythm, duration, and timbre, as well as a tonal memory task (Milner, 1962). Neither left nor right lesions affected performance on the pitch discrimination task, determining if two sine wave tones were presented at the same or different pitch. However, patients with right temporal lobe lesions were impaired on a variety of tasks, most notably tonal memory and timbre discrimination, but they were impaired to a lesser extent on loudness, rhythm, and duration discrimination tests as well. Patients with left temporal lobe lesions were not significantly impaired on any of the tests. Results remained the same whether the lesion encroached upon Heschl's gyrus, the cortical region containing primary auditory cortex, or not. This study implicated right temporal lobe auditory regions in a variety of basic auditory tasks relating to musical processing.

A more recent study tested similar patients with left or right anterior temporal lobe removals and a matched non-lesioned control group on two pitch perception tasks (Johnsrude et al., 2000). The first measured the threshold at which subjects could discriminate the pitch of two tones, while the second task measured the threshold for determining whether tone pairs increased or decreased in pitch. Neither patient group's threshold differed from controls on the first task, suggesting that simple pitch discrimination is performed subcortically, a result consistent with both human and animal literature (Meyer & Woolsey, 1952; Whitfield, 1980; Peretz et al., 1994). But for patients with right temporal lobe lesions that encroached upon primary auditory cortex, determining pitch direction caused great difficulty. They were able to perform the task, but their thresholds were significantly higher than those of controls. No other group was impaired, including patients with right temporal lobe lesions sparing primary auditory cortex. Therefore, while no impairments were seen on simple pitch discrimination with any unilateral temporal lobe lesion, the results suggest a role for right primary auditory cortex in higher order pitch processing.

Another study tested patients with unilateral lesions in temporal and/or frontal cortex on two pitch discrimination tasks (Zatorre & Samson, 1991). The first task required a simple pitch discrimination between two tones. As in the studies just described, all patient groups performed comparable to normal controls on this task. However, when interfering tones were placed between the two tones to be compared, thus increasing the task difficulty, patients with right temporal and/or frontal lesions were impaired. This study was important because it found evidence that in addition to the involvement of right temporal lobe auditory areas in pitch processing, right frontal regions are also involved when the task requires short term memory of tones.

A lesion study examining melodic discrimination and recognition emphasizes the importance of carefully analyzing this type of data with respect to exact lesion location (Zatorre, 1985). This study tested patients with right or left anterior temporal or right frontal lobe lesions and a group of matched controls. One of the tasks was to determine if two short novel melodies were exactly the same or not. On different-melody trials, one tone's pitch was altered such that the contour and key of the original melody were preserved, or the wrong note violated one or both of these features. The frontal lobe lesions did not affect performance on this task. The right temporal group was the only group impaired overall compared to controls, but as with each of the other groups, they were able to use both contour and key cues to improve performance. It was also found that when Heschl's gyrus was excised on either the right or the left, discrimination was impaired compared to controls. This impairment was additive with a right temporal lobe lesion such that a right temporal removal including Heschl's gyrus caused a greater impairment than a right temporal removal sparing this area, or a left temporal removal including this area. Only patients with left temporal lesions sparing Heschl's gyrus were unimpaired compared to controls. Therefore, this study suggested that detecting pitch changes within the context of a melody involves processing in both left and right primary auditory cortex as well as more anterior auditory regions on the right temporal lobe.

The importance of fractionating musical processes into distinct testable tasks is well illustrated by another two studies testing patients with cerebral lesions (Peretz, 1990; Liégeois-Chauvel et al., 1998). These two related studies tested patients with lesions in either the right or left temporal lobe on a variety of melodic discrimination tasks, with three tests varying different aspects of the pitch dimension, and two tests varying the temporal dimension. The first study tested patients with large cerebrovascular lesions in either the right or left hemisphere due to stroke (Peretz, 1990). Dissociations of function were observed such that different patterns of performance were produced by patients with lesions in the right or left hemisphere, with certain tasks disrupted by one, the other, or either hemisphere. For example, patients with right but not left hemisphere lesions were impaired using melodic contour as a discrimination cue, whereas patients with lesions in either hemisphere had difficulty discriminating melodies based on intervallic information alone. Interestingly, four patients exhibited a double dissociation of pitch and rhythm perception such that performance on pitch tasks was intact while rhythmic perception was impaired, or vice versa. A second study tested a series of patients with more focal lesions in the right or left temporal lobe on the same tasks (Liégeois-Chauvel et al., 1998), replicating the results of the previous study and refining the anatomical localizations of particular musical functions.

These studies suggested that pitch processing in the form of melodic discrimination is performed hierarchically by separate neural systems, such that the perception of intervals is affected by processing in both right and left temporal lobes, whereas contour perception is affected only by right temporal lobe processing. However, a previous study investigating the use of contour information in melodic discrimination found that while right temporal lobe lesions affect melodic discrimination more severely, patients with these lesions were still able to use contour information to the same extent as controls or patients with left hemisphere lesions (Zatorre, 1985). This discrepancy may reflect differences in exact lesion locations, the specific tasks tested, or way these tasks were compared. But despite the apparent differences in specific results, the complexity of interactions between task and lesion location in these studies highlights the idea that music perception is a conglomerate of a variety of different perceptual processes, each treated differently by the brain.

Functional neuroimaging

Current functional neuroimaging methodologies such as positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) (summarized below) allow researchers to peer inside normal living brains and see which neural structures are active when someone performs a specific cognitive or perceptual task. These studies generally support the lesion literature by demonstrating right hemispheric structures as being involved in tasks requiring pitch or melodic processing, particularly right temporal and frontal regions (Zatorre et al., 1992; Zatorre et al., 1994; Binder et al., 1997; Holcomb et al., 1998; Griffiths et al., 1999; Halpern & Zatorre, 1999; Perry et al., 1999). In one of the first PET studies to include a pitch judgment task, right frontal regions were found to be more active when judging the direction of pitch change in speech syllables than passive listening to those syllables (Zatorre et al., 1992). Making phonetic judgments on the same syllables did not activate these regions, suggesting that distinct neural areas are specifically involved in pitch processing.

A PET study investigating melodic perception found the right superior temporal gyrus (STG) to be preferentially active during passive listening to melodies (Zatorre et al., 1994). This activity mirrors the lesion studies that find pitch processing can be disrupted by lesions in right temporal lobe regions (Milner, 1962; Samson & Zatorre, 1991; Zatorre & Halpern, 1993; Johnsrude et al., 2000). When a pitch judgment of the first two notes was required, right frontal regions also became active, again supporting lesion studies that find higher level pitch processing to be disrupted by lesions in the right temporal and frontal regions (Zatorre & Samson, 1991). By requiring subjects to compare the pitch of the first and last tones of each melody, increasing the memory load and thus task difficulty, a wider pool of cortical areas was recruited, including right frontal, temporal, and parietal regions, with both hemispheres represented, but more activity evident on the right.

Right frontal activations seen in the functional neuroimaging of pitch discrimination tasks are generally interpreted as being involved in the upkeep of tones in working memory while making the comparisons. Frontal regions in both hemispheres have been implicated in various working memory tasks, (Owen, 1997; Petrides, 1998), and right frontal regions in particular have been shown to be involved in tonal working memory (Zatorre & Samson, 1991).

The neuropsychological literature on tonal and melodic processing has progressed in leaps and bounds, particularly in the past decade or so. The benefits of studying the neural substrates of cognitive processes through these types of stimuli have become widely known and accepted, creating a fertile ground in the study of cortical function. Through careful manipulation of both stimulus, task, and type of participant tested (e.g. musicians, nonmusicians, specific patient populations, etc.), a comprehensive understanding of the neural processes involved in perceiving music specifically, and sounds in general, will be reached.

Complementarity of different methodologies

The four projects included in this thesis utilized four different, complementary methodologies: 1) Behavioral measures in the normal population, 2) Behavioral measures in people with focal brain excisions, 3) Functional neuroimaging with positron emission tomography (PET), and 4) Functional neuroimaging with functional magnetic resonance imaging (fMRI). Each of these methods allowed examination of the research question from a different perspective. As the advantages and disadvantages of these methods differ, it was possible to use the different methodologies to focus on different aspects of the question, developing a wider view of the process than would be possible with one or two methodologies alone. A brief overview of each method including some of their advantages and disadvantages follows.

Behavioral testing of a normal population

Before beginning an investigation into the neural correlates of a cognitive process, it is first necessary to understand the nature of that process in a normal working brain. This is generally done by observing how the process manifests itself behaviorally in a normal population. When the research question pertains to the neural correlates of a perceptual phenomenon, as in this thesis, it is first necessary to establish a) that the phenomenon exists, b) insights into its nature, and c) an efficient way to test it behaviorally. A behavioral study of the process in question can address these issues, as well as generate a quantitative behavioral baseline, laying the groundwork for subsequent studies. Although the behavioral methodology is very important in defining the nature of a perceptual phenomenon, it does not lend any clues as to its neural correlates. Other methodologies are needed for addressing this question.

Behavioral testing of patients with cortical lesions

The classic methodology for determining neural substrates of a particular cognitive process involves testing patients with known brain damage on behavioral tasks.

The key to this method is knowing the location of each patient's lesion. With this knowledge, one can look for consistent task deficits in patients with similar lesions. If this pattern is found, it is possible to postulate that the common damaged cortical region is involved in the neural processing necessary for carrying out that task. However, in human patients, this ideal is often difficult to obtain. One difficulty is that human lesions, created either by naturally occurring circumstances or individually tailored neurosurgery, are seldom uniform. Therefore, determining the exact location of these lesions and carefully analyzing the data according to lesion site is necessary for gaining useful information from these studies.

Although the use of modern anatomical imaging methods can determine precise lesion locations, behavioral testing of patients with cortical lesions does not necessarily pinpoint regions functionally involved in a particular task as focally as functional neuroimaging methods. Because the lesion overlap between patients can be large, it may not be possible to tell if the region contributing to observed deficits comprises of the entire overlapping area, or a smaller region within that area. In addition, once a region is isolated as being involved in task processing, it is not possible to discern from that study whether the area is sufficient for that processing, or if it is part of a larger system of contributing structures. However, the converse of this argument is advantageous to the lesion study method. For example, in the clearest case of lesion study results, patients with no damage to a certain structure perform a task as well as a control group, while patients with damage including that area cannot perform the task. In this case it is possible to assume that the implicated structure is necessary for task performance, a claim that cannot be made in functional neuroimaging studies.

Another difficulty with this methodology, particularly when testing epilepsy patients as was done for this thesis, stems from inferring normal cerebral function from cortex afflicted with a chronic neurological problem. It is not prudent to assume that the disease necessitating a portion of cortex to be resected has not affected the remaining tissue, particularly if the problem began early in life, as it is possible in these cases that unafflicted cortex went through some amount of functional re-organization.

Caveats aside, testing patients with known cortical lesions is a powerful method for determining neural correlates of a particular cognitive process. It is possible to use this method to isolate particular structures as being involved in the processing of a particular task. In addition, these structures can then be inferred to be *necessary* for the task processing, a claim that can not be made based on functional neuroimaging results.

Functional neuroimaging with Positron Emission Tomography (PET)

PET is another method used to locate cortical regions involved in different cognitive and perceptual tasks. The spatial localization obtainable with this method is much finer than lesion studies, but less precise than fMRI. It involves injecting a subject with a low dose of radioactive tracer, usually incorporated into water, which travels through the bloodstream to the brain (Fox et al., 1984). When the subject performs a task, the cortical areas involved in that task draw more blood flow, and therefore more tracer, than other areas. As the tracer decays, it emits positrons. When a positron hits an electron, two gamma rays are emitted in diametrically opposite directions. The PET scanner detects each pair of gamma rays and backprojects an image from the detectors to where in space (and therefore the brain) the gamma rays were emitted. This image is how the PET data is viewed. By co-registering the functional data from PET scans with anatomical magnetic resonance imaging scans, it is possible to determine more precisely the anatomical regions of activity (Evans et al., 1991; Watson et al., 1993).

One advantage of PET over lesion studies is that it allows us to peer inside a normal working brain, and thus avoid any problems associated with neural processes of damaged brains. In addition, this method allows much more precise anatomical localization than the patient lesion method. Also, whereas lesion studies require choosing a single or small number of cortical regions to investigate, the entire brain is imaged by the PET scanner, enabling whole brain exploratory searches of the data. This allows researchers to observe the systems of structures involved in a task, and lifts the restriction of looking at only one or two regions at a time.

Despite the advantages of PET over lesion studies, there are also some disadvantages. As described above, although functional imaging techniques illuminate systems of structures involved in a particular task, these methods cannot determine which, if any, of those structures are *necessary* for the task. In addition, a PET scan is a somewhat invasive technique; subjects are injected with low-level radioactive material.

Another difficulty with this methodology is that each scan lasts only one minute, so the number of trials presented is quite limited, even when repeating the same condition over a couple of scans. The number of scans allowed per subject is also limited due to dosimetry considerations.

Functional neuroimaging with Functional Magnetic Resonance Imaging (fMRI)

fMRI is another functional neuroimaging technique used for imaging normal living brains. Without the need for injections or radioactivity, it uses series of magnetic pulses to detect changes in blood oxygenation level associated with haemodynamic changes correlated with neuronal activity (Moseley & Glover, 1995). As in PET, it relies on the idea that active cortical regions will recruit more blood than those at baseline levels of activity, but fMRI allows finer spatial localization than PET. Another advantage of fMRI over PET is that since it is not restricted to one minute scanning blocks like the PET paradigm, it is possible to include many more trials, increasing signal to noise ratios.

Different types of paradigms can be used with this methodology, depending on the nature of the research question. The traditional block format (necessary for PET studies) can be employed, in which all stimuli of a particular type are grouped together and scans are taken periodically during each block (Belliveau et al., 1991). Alternately, as in an evoked potential paradigm in electroencephalography (EEG), the stimuli can be randomized together with one scan per stimulus, timing the scan to record maximum haemodynamic response to a particular aspect of each stimulus (Buckner et al., 1996). The fMRI study reported in this thesis uses this second paradigm. It is also possible to track the time-course of the haemodynamic response to a particular stimulus within specified cortical structures by a process that varies the time of the scan relative to stimulus presentation (Belin et al., 1999).

Although it is still far from the millisecond timing possible with EEG, fMRI shows a marked improvement in timing resolution as compared to PET, allowing timing in the realm of a few seconds, regulated by the speed of the haemodynamic response, while PET averages all activity over the course of a minute-long scan. This improvement in timing accuracy allows researchers to examine "event-related" neuronal activity as opposed to the "state-related" activity imposed by timing considerations of PET (Friston, 1997).

However, fMRI has its own drawbacks. For instance, like PET, it is not possible to tell which of the observed activations are necessary for task processing, something only possible with lesion methods. Also, the small enclosed space into which the subject must be placed, combined with the magnetized environment makes the use of stimulus presentation and auxiliary measurement apparatus such as headphones, heart rate monitors or EEG equipment difficult. The scanner itself is also a very noisy machine, making fine auditory stimulus presentations difficult. Fortunately, a team of researchers has found a way to get around the noise problem by spacing stimuli so that the haemodynamic response to scanner noise will have died down before taking a scan related to the intended stimulus (Belin et al., 1999). The increased resolution in both spatial location and timing resolution as compared to PET, combined with relatively few disadvantages, make the fMRI a useful tool for localizing neural correlates of cognitive processes.

Importance of complementary and converging findings

The type of information obtained from each of the four methods described above is different, yet complementary. Different assumptions are associated with each methodology, and similar results across differing methodologies help to confirm that the results are due to a real phenomenon, and are not the product of a problematic assumption. Therefore, complementary or converging findings allow the researcher more confidence when interpreting conclusions from the results.

In addition, some questions are better answered with one method than another, and the combination of results allows a much wider view of the situation than results from one method alone could produce. Normal cognitive processes are best delineated by behavioral testing of normal subjects, but except in rare cases such as dichotic listening paradigms, this method does not divulge any information related to localization of function. Functional localization requires either behavioral testing of brain lesioned patients, or functional neuroimaging techniques. These localization methods also work together in tandem. For example, a PET study could help to determine if a structure implicated in a lesion study was part of a larger network of cortical regions, and conversely, a lesion study could help determine whether a structure seen in an fMRI study was necessary to perform the task or not. An fMRI study could potentially determine the time course of activation of structures seen in a PET study. EEG, a methodology attempted but not included in this thesis, could further illuminate this time course to within millisecond accuracy. In short, the strengths of different methodologies allow questions to be asked in different ways, and as more results from various studies and methodologies converge, stronger statements can be made about the answers.

Brief overview of thesis projects

Behavioral testing of normal population: Chapter 2

The first study included in this thesis is a behavioral study of the normal population. The perceptual question asked was: How do timbre and tonal context affect pitch perception? More specifically, what is the effect of spectral timbre on pitch perception, and how does tonal context affect it? There were many clues in the cognitive and perceptual literature regarding the effect of context or timbre on pitch separately, but they had never been combined to create a larger picture.

The paradigm used combined pitch and timbre by varying them separately and together in different stimuli, creating a difficult pitch constancy task. The tonal context surrounding the stimuli were also varied in order to determine if context could make the task easier. This study confirmed that spectral timbre strongly affects pitch perception when discriminating small differences in fundamental frequency. In addition, this task was made much easier when it was placed within the context of a melody, thus confirming the existence of the perceptual phenomenon we wished to test, as well as how it manifested itself behaviorally.

Since the paradigm would eventually be used for testing patients, it was constructed so that a significant effect could be measured in a manner suitable for patient testing. The test therefore took relatively little time to administer, as time with the patients is very valuable, and the task was not too complex, so that patients with lower IQs would not be at a disadvantage. Musically untrained subjects were tested so that comparisons could be made between these results and those obtained with a patient population.

Behavioral testing of patients with cortical lesions: Chapter 3

The next step in this series of studies was to apply the paradigm used in Chapter 2 to a group of epilepsy patients who had received unilateral cortical excisions of anterior temporal lobe as a treatment for the seizure disorder. This study was the first step towards isolating where within the auditory system neural processing related to the context effect seen in the previous behavioral study could be occurring. It was possible to determine that auditory association areas on the anterior portion of the right temporal lobe must be intact in order to obtain the facilitating effects of melodic context on pitch perception.

Functional neuroimaging with fMRI: Chapter 4

Using the patient study results to guide predictions, an fMRI study examining tonal context effects was performed on a group of normal subjects. Since neither timbre nor context effects on pitch perception had been studied with neuroimaging techniques, the paradigm was simplified to avoid difficulties when interpreting the results, varying only the tonal context and keeping timbre constant across all stimuli. Both behavioral and functional data were combined in this study, allowing results from one set of data to aid in the interpretation of the other. Although the paradigm in the fMRI study differed slightly from the behavioral studies discussed above, the results confirmed those found in the patient study, and allowed more precise anatomical localization of function. This convergence of results strengthened the interpretation of each individual study.

Functional neuroimaging with PET: Chapter 5

The PET study included in this thesis does not follow directly from the studies discussed above, but deals with some of the same issues, and expands the scope of the thesis to include a comparison between tonal and linguistic processing. The stimulus used in this paradigm, sung words, combines tonal and linguistic aspects of sound into a single percept, allowing a comparison of pitch and semantic judgments of the same stimuli. The semantic judgments are made relative to the linguistic content of the lyric, and, similar to chapters 2-4, pitch judgments are made within a melodic context. By comparing PET activations related to melodic and semantic monitoring, we were able to dissociate regions of activity specific to each of these stimulus attributes. While some regions were activated by both melodic and semantic monitoring, other regions were either more active or only active in one of the two conditions, suggesting specialization of function in these regions.

Conclusion

This introductory chapter serves as an overview of the general topics addressed in the remainder of this thesis. Some of these topics are touched upon again in the introductions of individual studies in the next four chapters, but due to length limitations when submitting these studies for publication, it was not possible to go into as much detail as has been addressed in this chapter. Similarly, the length restrictions limit the amount of detail included in the discussion sections of these chapters. Therefore, a more comprehensive discussion of each of the results, as well as how they inter-relate to each other and to the relevant literature is included in the last chapter of the thesis entitled General Discussion.

Chapter 2

Influence of Tonal Context and Timbral Variation on Perception of Pitch

Pitch and timbre are two of the building blocks of music. Variations in pitch lead to a melodic line, while variations of timbre are usually heard as different instrumentations. Do pitch and timbre interact? Listening to an orchestra, one can hear a continuous melody being played when different instruments switch off playing separate parts of this melody, even if each note of the melody is played by a different instrument, a compositional style called klangfarbenmelodie or hocket. A more striking demonstration of this phenomenon occurs with sung melodies, in which the changing vocal timbres associated with speech do not alter perception of the melody. This implies that one can follow the fundamental frequency (F0) of a series of tones, even when their spectral shapes differ, arguing for the separability of pitch and timbre. However, since both pitch and spectral timbre are rooted in the frequency dimension of sound, it should not be surprising if they interact under some circumstances. To test this idea, one could look at people's perceptions of pitch and/or timbre when both the F0 and the spectral shape of the tones differ. This paper investigates the interaction between pitch and timbre, focusing on spectral timbre's influence on pitch perception as a function of context.

The literature examining interactions between pitch and timbre has yielded contradictory results. Some researchers have found the timbre of a tone to affect its perceived pitch (e.g. Singh and Hirsh, 1992; Platt and Racine, 1985; Krumhansl and Iverson, 1992, exp 1; Melara and Marks, 1990a,b,c; Wapnick and Freeman, 1980), while others have found no effect of timbre on pitch perception (e.g. Semal and Demany 1991, 1993; Krumhansl and Iverson, 1992, exp. 2,3). It seems that those studies presenting tones in the absence of other tones tend to find an interaction between pitch and timbre, while studies presenting tones within the context of other tones find no such interaction (but see also Demany and Semal, 1993, in which pitch and timbre difference thresholds for isolated tones were not affected by variation in the irrelevant dimension).

A same-different paradigm was used by Singh and Hirsh (1992) to determine the perceived pitch of isolated residue tones: tones having no component at F0. Six timbres were synthesized, each containing four consecutive harmonics differing only on the lowest harmonic. Each pair of tones could differ in F0, spectral composition, or both. Participants indicated whether timbre was the same or different, and whether pitch stayed the same, went up, or went down. When the harmonic numbers and F0 moved in the same direction, participants correctly reported the direction of pitch change. However, when harmonic number and F0 moved in opposite directions, this created a conflict. If the change in F0 was less than 4%, the direction of harmonic change dominated, and when the change in F0 was 4% or greater, the pitch direction dominated. Therefore, pitch and timbre were found to be separable only with a change in F0 of 4% or higher.

Platt and Racine (1985) used a tuning paradigm to examine differences in the perceived pitch of pure and complex tones. Larger tuning deviations were made tuning a complex tone to a pure tone than when both the standard and test tones were pure tones, suggesting an interaction between pitch and timbre. Classifying single tones into categories, another paradigm performed without a tonal context, also finds pitch and timbre to interact (Krumhansl and Iverson, 1992, exp 1; Melara and Marks, 1990a,b,c). Melara and Marks performed a series of experiments investigating pitch, timbre, and loudness interactions using the Garner speeded classification method (Garner, 1974), observing how redundant and interfering information from the unattended dimension affects reaction times, and looking at classification reaction times to selective and divided attention to these dimensions with varying orientations of stimulus axes. These studies illustrate a consistent interaction between pitch and timbre; correlating information from the unattended dimension enhances classification, while competing information disrupts it. They also determined that both pitch and timbre are "hard," "primary" perceptual dimensions (Melara and Marks, 1990a,b,c). Krumhansl and Iverson replicated the Garner classification results with different stimuli, finding pitch and timbre to interact in isolation. They then went on to test for these interactions within longer sequences, and found a much different result; changes in pitch of the context tones affected the pitch but not timbre perception of the test tones, and changes in timbre of the context tones weakly affected their timbre but not pitch perception (Krumhansl and Iverson, 1992).

This lack of interaction is seen in other studies examining pitch and timbre in a tonal context. Semal and Demany performed a series of experiments finding no interaction between pitch and timbre (1991, 1993). In both studies, they presented eight-tone sequences and asked participants to judge whether the last tone was identical or different than the first tone. The six interpolated tones varied by F0, spectral timbre, amplitude envelope, and intensity. However, only the F0 of the interference tones affected performance; timbral variations had no effect. Therefore, it was argued that pitch and timbre are completely separable. In a study using a similar paradigm, variations in pitch were not found to significantly affect timbre perception (Starr and Pitt, 1997).

It appears that interactions between pitch and timbre tend to occur in situations in which test tones are presented in the absence of other tones. Methodologies which present tones in the absence of a tonal context find pitch and timbre to influence each other: comparing a single tone to another single tone by tuning (Platt and Racine, 1985), making qualitative judgments (Sing and Hirsh 1992; Wapnick and Freeman, 1980), or classifying single tones into categories (Krumhansl and Iverson, exp. 1, 1992; Melara and Marks, 1990a,b,c). In contrast, those studies which present tones in the context of other tones find no evidence of an interaction: manipulating interpolating tones (Semal and Demany, 1991, 1993; Starr and Pitt, 1997), or comparing the fourth tones of two seven-tone sequences, a modified interpolated tone paradigm (Krumhansl and Iverson, 1992, exp.2,3).

The fact that tonal context can affect pitch perception has been well established (Deutsch 1972a,b, and 1982; Deutsch and Roll 1974; Dewar et al, 1977; Krumhansl 1979; Krumhansl and Castellano, 1983). After performing a series of studies on the matter, Krumhansl concluded that "the representation of pitch consists of a pattern of interrelationships that is highly specific to the tonal system of the musical context. Further, this pattern of interrelationships, once established, has implications for the processing of subsequent musical events" (1979, p.372). The present study tests whether or not tonal context's effect on pitch can also affect the interaction between pitch and timbre, possibly explaining the differences seen in previous studies. Our goal was to examine the effect of context specifically, while controlling for effects of interpolating tones by utilizing a methodology that avoided the use of any interpolating tones. Two contexts were used in experiment 1: 1) an isolated context in which two tones only were presented for each trial, and 2) a melodic context in which these same tones were presented as the final tone of a familiar melody. The task in each condition was to make pitch judgments while ignoring changes in timbre. We hypothesized that the presence of other tones in the melodic context would provide more of a tonal reference point from which to judge pitch so that people would make more accurate pitch judgments and be better able to ignore differences in spectral composition in the melodic context than in the isolated context.

EXPERIMENT 1

Method

Subjects: Eleven McGill undergraduates participated in this experiment. None had any extensive musical training (average training = 1.25 years). Normal hearing was determined by self-report. All were compensated for their time and attention.

Stimuli: The stimuli were digitally synthesized on an IBM compatible 386 computer using MITSYN software (Henke, 1981). All test tones were 500 msec duration, with rise and fall times of 10 msec. Contextual tones used in the melodic condition varied in duration according to the score. Three timbres were created by varying the relative intensities of 11 harmonics, keeping the intensity of the fundamental frequency constant. One emphasized the lower harmonics (Low), another emphasized the middle harmonics (Middle) and sounded brighter than Low, and the last emphasized the higher harmonics (High) and sounded brighter than Middle (see Figure 1). F0 was kept at a constant level in each timbre in order to avoid basing differences between tones on F0 strength. Individual tones ranged in fundamental frequency from 164.81 – 1108.70 Hz. (E3 to C#6), with test tones always presented between 261.63 Hz (C4) and 480.35 Hz (A#4 + 52¢), and all were equated for loudness. Stimuli were presented with MAPLE (Manager of Auditory Perception and Linguistic Experiments) software (Achim et al, 1992), and run through a passive Tchebyshev function lowpass filter having a 3dB cutoff at 8000 Hertz, and a slope of –142dB per octave. Sounds were presented binaurally at 75 dB SPL over Seinheisser HD424 headphones in a soundproof booth.

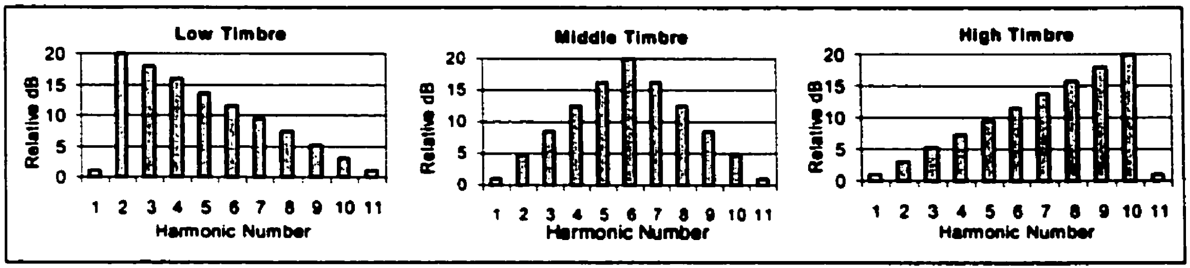

Figure 1: Spectral shapes of timbres. Differences in intensity level of each harmonic are shown relative to the other harmonics.

Procedure

Isolated context: The concepts of pitch and timbre were explained, with timbre being explained as a change in "instrument" or "sound quality," and exemplars of the three spectral shapes were presented at a common F0. Each trial consisted of two tones presented with an interstimulus interval (ISI) of 100 msec. The first tone was randomly presented at one of six frequencies corresponding to whole tone steps starting at C4 (261.63 Hz) and continuing to A#4 (466.16). The second tone was 0 (same), 17, 35, or 52¢ higher than the first tone, corresponding to the 1, 2, and 3% differences at which Singh and Hirsh found pitch and timbre to interact. (Cents are logarithmically equal steps in the frequency dimension, each semitone being 100¢ apart.) Each pair of test tones could be presented in one of three timbre pairings: 1) same (Low-Low, High-High), 2) small difference (Low-Middle, High-Middle), and 3) large difference (High-Low, Low-High). Each pairing was presented at all levels of F0 difference. Participants were informed that they would be making pitch judgments on these tones, and that they should ignore any timbre changes as much as possible. When the tones were presented, participants were instructed to indicate whether the pitch of the second tone was the same as or different from the first tone by pressing a key on the computer keyboard. If they responded "different," they were asked to indicate how different on a scale of 1-3 where 1 = slightly different and 3 = very different. Subjects completed 36 trials at their own pace after completing a block of 6 practice trials.

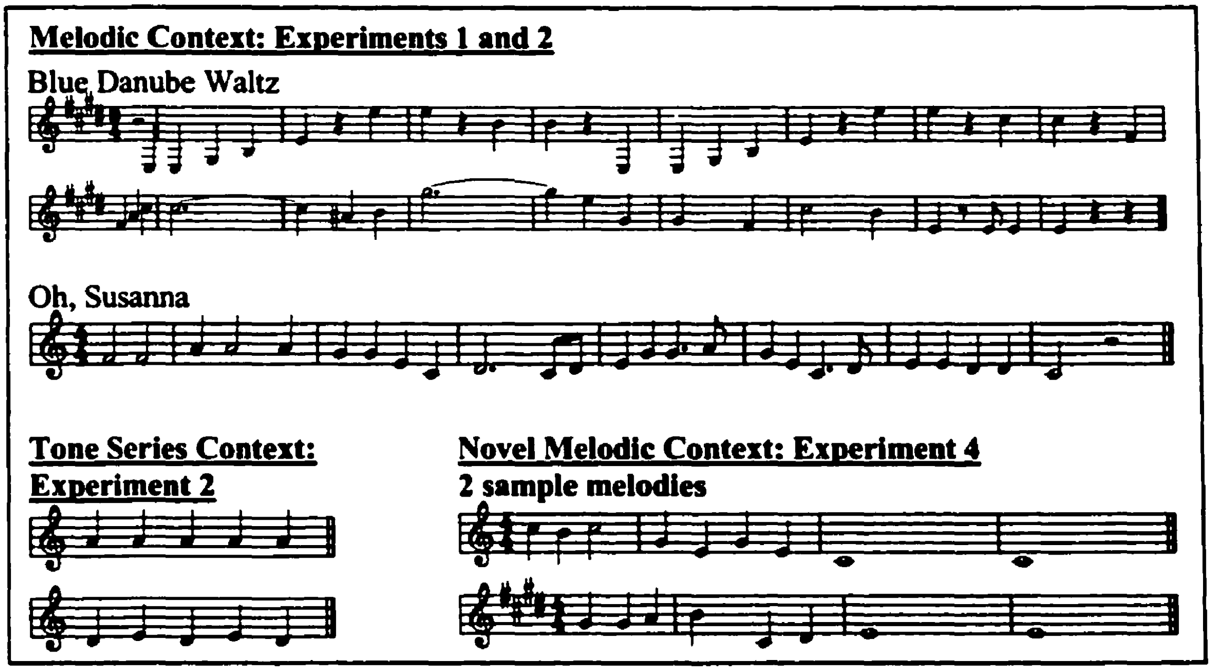

Figure 2: Stimulus contexts shown in musical notation.

Melodic context: For the melodic context, the task was to determine whether the last note of a melody was in tune or not. One of two melodies, familiar to all participants, was presented on each trial: "Oh, Susanna," or "The Blue Danube Waltz" (See figure 2). These melodies were chosen for their different endings; the four final notes of "Blue Danube" all share the same pitch, while "Oh, Susanna" ends in a downward scalar motion. Each melody, excepting the last note, was presented in tune in one of three keys so that the last note should either end on one of six whole steps from C4 to A#4, and was presented in one of the three timbres described above. The last note could continue with the same spectral shape as the melody, or sound in one of the other two, thus creating the same three timbre pairs as the isolated condition, and could be 0 (same), 17, 35, or 52¢ sharper than the correct final note. After each melody, participants were instructed to indicate if the last note of the melody was in or out of tune, ignoring any timbre differences. If they responded "Out of Tune," they were asked to indicate how much out of tune on a scale of 1 - 3 where 1 = slightly out of tune, and 3 = very out of tune. Subjects completed 36 trials of each melody at their own pace, after a block of 6 practice trials. We predicted that if there was a difference in responses between the two melodies, subjects would do better on “Blue Danube” trials because they could compare the pitch of the last note directly to that of the penultimate note.

Results

Results for each trial were coded on a scale of 0-3 where 0 indicated a “Same” response, and 1-3 corresponded to the scales of 1-3 described above. Analyses of variance (ANOVA’s) were performed within subjects as repeated measures tests with three factors: context (isolated vs. melodic), F0 (0, 17, 35, or 52¢), and timbre-pair type (same, small difference, or large difference). Tukey’s HSD test (Q statistic) was used for post-hoc analyses.

In the melodic condition, responses to both melodies were similar, with increasing responses to larger timbre differences and larger F0 deviations (See figure 3). An interaction between timbre-pair-type and melody was found, so we analyzed responses to each timbre pair type across melodies. In this analysis, F0 deviations in different-timbre trials in “Blue Danube” were judged larger than in “Oh, Susanna” $F(1,10) = 17.21$, $P<0.01$ for small timbre difference, and $F(1,10) = 10.11$, $P = 0.01$ for large timbre difference). When collapsing across timbre pair types, subjects judged F0 differences consistently across melodies. Since both melodies differed significantly from the isolated context in the same direction and manner, while differing much less between themselves, data from the two melodies were averaged together and called the “melodic condition” in further analyses.

Responses to the isolated condition also increased with increasing differences in timbre. But unlike the melodic condition, responses did not increase with increasing F0 deviations (See figure 3).

Statistical comparisons between the two contexts showed a main effect of timbre pair type $F(2,20) = 131.22$, $P<0.001$ and an interaction between context and F0 $F(3,30) = 19.68$, $P<0.001$. In investigating this interaction, we conducted two types of post-hoc analyses. First, we compared the ratings of each F0 level between contexts, collapsing across timbre pairs. No differences were found between contexts in either the 0 or 17¢ deviations $F(1,10) = 1.04$ and 1.47 respectively; $P>0.2$ for both. Subjects rated 35 and 52¢ trials as more different/out of tune in the melodic than in the isolated context,

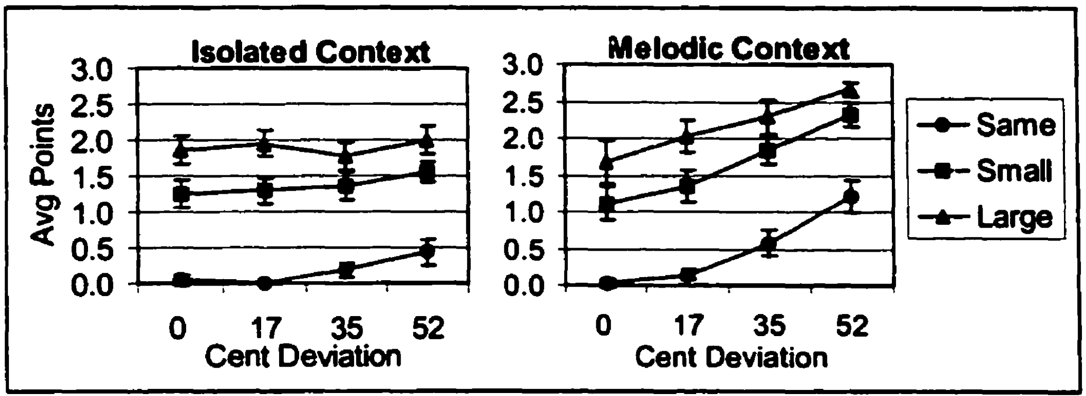

Figure 3: Experiment 1 results averaged across subjects. Scored on scale of 0-3 where 0 means "Same Pitch" (See text for details). In timbre differences legend, "Same" denotes same timbre, "Small" denotes a small difference in timbre, and "Large" denotes a large difference in timbre. Standard error bars shown.

indicating better discriminability of F0 in the melodic condition $Q(8,30) = 7.7$, $P<0.05$ for 35¢; $Q(8,30) = 12.4$, $P<0.05$ for 52¢.

We also investigated which of the F0 differences were detected within each context by comparing ratings on "different F0" trials to those of the "same F0" trials, again collapsing over timbre pair types. The only F0 difference detected in the isolated condition was the 52¢ deviation $Q(4,30) = 5.22$, $P<0.05$. In contrast, both the 35 and 52¢ deviations were detected in the melodic condition, again arguing for better F0 discrimination in the melodic condition $Q(4,30) = 7.71$, $P<0.05$ for 35¢; $Q(4,30) = 13.88$, $P<0.05$ for 52¢.

Discussion - Experiment 1

The most salient effect seen in this experiment was the influence spectral shape had on the detection of small changes in F0 in both the isolated and melodic conditions. At all levels of F0, tones that differed in timbre were judged more different in pitch than when timbre was the same. This replicates the results found in Singh and Hirsh's 1992 study in which timbre differences interfered with detection of F0 deviations of a similar magnitude. However, this experiment also showed that placing tones within a melodic context allowed listeners to perceive F0 differences that were undetectable without this contextual information. This facilitation occurred despite the continuing strong effect of timbre on pitch judgments.

This facilitation of pitch discrimination within a melodic context could be due to the fact that the melodies are creating a tonal reference point to which the test tones can be related. According to Deutsch, "in listening to sequences, we process not only the individual tones, but also the melodic intervals between them. These intervals then provide a framework of pitch relationships to which the test tones can be anchored" (1982 p. 302). The tonal structure of the melodic condition, unavailable in the isolated condition, helps us to focus in on the pitch information carried in the tones, and thus seems to facilitate the separation of timbre from the pitch percept.

While addressing some of the issues regarding the interactions between pitch, timbre, and context, experiment 1 also raises new questions. For example, exactly what aspect of the melodic context allows pitch to be perceived more accurately than when this context is not present? Would the simple presence of other tones be enough to show this effect, or is the tonal structure of the melodies necessary?

A second experiment therefore included a third context in which the same tones were placed at the end of non-melodic five-tone sequences which mimicked the endings of the melodies used in the melodic condition of experiment 1. We predicted that if the effect of context seen in experiment 1 was caused by the simple presence of other tones, results of the tone-series condition would be no different than the melodic condition. However, if the tonal structure of the notes is necessary to see an effect of context, the tone-series condition would yield results more like the isolated condition, and if the effect was due to both the tones' presence and their tonal structure, performance on the tone-series condition would lie between the isolated and melodic conditions due to the presence of one but not the other element.

EXPERIMENT 2

Methods

Subjects and Stimuli: Twelve McGill undergraduates participated in this experiment. None had any extensive musical training (average length = 1.67 years). No participants reported any hearing loss. All were compensated for their time and attention. Stimuli were synthesized and presented as in experiment 1.

Procedure: The isolated and melodic contexts were presented as in experiment 1, with the exception that the 17¢ deviation of F0 was dropped because it was indistinguishable from the 0¢ difference in both conditions. In addition to these two contexts, experiment 2 contained a third context in which the test tones were placed at the end of two simple five-tone sequences. This context was presented in the same manner as the melodic condition, except that instead of hearing melodies, participants heard tone sequences. One sequence simply repeated one tone five times, while the other alternated between one tone and a whole tone above it, mirroring the endings heard in the melodic condition (See figure 2). The order of the three conditions was counterbalanced across subjects to control for order effects.

Results

As in experiment 1, results for each trial were coded on a scale of 0-3. Analyses of variance were performed within subjects as repeated measures tests with three factors: context (isolated, tone series, and melodic), F0 (0, 35, and 52¢), and timbre-pair type (same, small difference, and large difference). Tukey HSD tests were performed for post-hoc analyses.

The pattern of results in both the isolated and melodic contexts replicated experiment one's results (See figure 4). Responses to both melodies again increased with both increasing timbre and F0 difference. F0 deviations in different-timbre trials were again judged larger than in "Oh, Susanna" $F(1,11) = 16.02$, $P<0.01$ for small spectral differences, $F(1,11) = 23.08$, $P=0.001$ for large spectral differences, but the differences between melodies were much smaller than their mutual differences from the isolated condition, and so the two melodies were grouped together for further analysis as the melodic condition. No consistent differences were found between responses in the alternating and repeating tone conditions, so they were also averaged together for further analysis as the tone-series condition $F(1,11) = 3.91$, $P>0.05$.

Statistical comparisons between contexts uncovered a three-way interaction between context, timbre-pair type, and F0 $F(8,88) = 3.83$, $P<0.01$. In investigating this interaction, we again conducted two types of post-hoc analyses, both using Tukey's HSD test. First, we compared the ratings of each F0 level between contexts at each of the three spectral pairings. The pattern of results indicated an increased discriminability of F0 with increasing levels of context.

Most of the differences found were seen when comparing the isolated to the melodic context. When spectral shape did not differ, both the 35 and 52¢ deviations were judged larger in the melodic context $Q(9,88) = 6.74$, $P<0.05$ for 35¢; $Q(9,88) = 9.38$, $P<0.05$ for 52¢. With small spectral differences, 0¢ deviations were judged smaller, and 35 and 52¢ deviations were judged larger in the melodic condition $Q(9,88) = 6.01$, $P<0.05$ for 0¢; $Q(9,88) = 5.29$, $P<0.05$ for 35¢; $Q(9,88) = 8.78$, $P<0.05$ for 52¢. With large spectral differences, the 52¢ deviation was judged larger in the melodic context $Q(9,88) = 7.34$, $P<0.05$.

Because the tone series results fall between those of the isolated and melodic contexts, they differ only slightly from either one. Comparing the tone series and melodic contexts, with small spectral differences, the 35¢ deviation was judged larger in the melodic context $Q(9,88) = 8.78$, $P<0.05$. With large spectral differences, the 52¢ deviation was judged larger in the melodic condition $Q(9,88) = 6.74$, $P<0.05$. The only difference seen between the isolated and tone series contexts was that people judged same F0 trials in the small timbre difference as smaller in the tone series context $Q(9,88) = 5.29$, $P<0.05$.

We also investigated which F0 differences were detected by comparing "different F0" trial ratings to "same F0" trial ratings within each timbre difference for each context; e.g. within large timbre difference trials, was the 35¢ deviation judged different than same F0 trials. None of the F0 differences were detected in the isolated condition. Three of the six F0 deviations were detected in the tone-series condition. The 52¢ difference was detected in both the same and small spectral difference trials $Q(9,88) = 7.70$, $P<0.05$ for same; $Q(9,88) = 9.26$, $P<0.05$ for small, and the 35¢ deviation was detected in the large spectral difference trials $Q(9,88) = 6.37$, $P<0.05$. All F0 deviations were detected in the melodic condition. This analysis also indicates increasing discriminability of F0 with increasing levels of context.

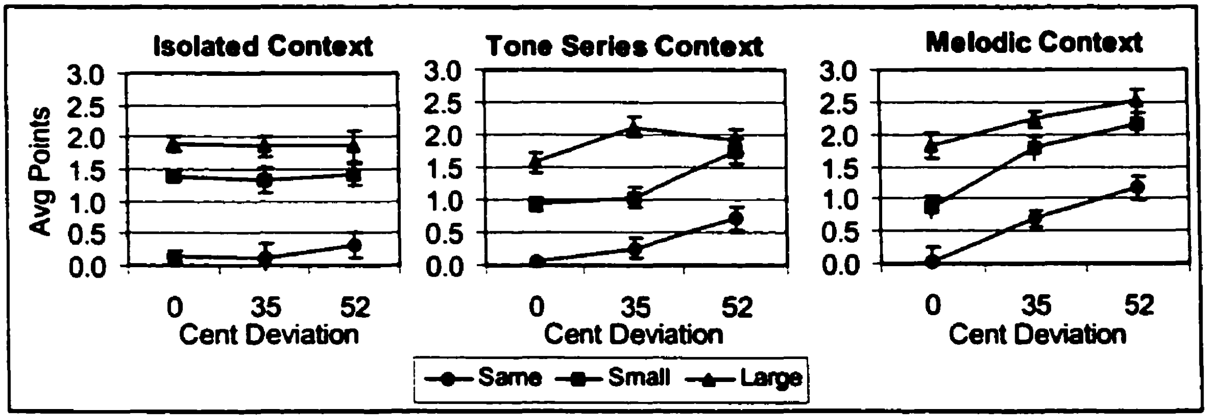

Figure 4: Experiment 2 results averaged across subjects. Scored on scale of 0-3 where 0 means "Same Pitch" (See text for details). In timbre differences legend, "Same" denotes same timbre, "Small" denotes a small difference in timbre, and "Large" denotes a large difference in timbre. Standard error bars shown.

Discussion - Experiment 2

The isolated and melodic conditions in Experiment 2 replicated experiment 1. The effect of spectral shape on pitch perception was again the most salient effect, being highly significant in all three contexts. Thus, participants were more likely to judge pitch as more different when timbre differed than when it was the same, regardless of differences in F0. When comparing the isolated and melodic conditions, we saw the same pattern of results found in experiment 1. Namely, despite timbre's influence on pitch perception, F0 differences that were undetectable in the isolated context were perceived when heard within the context of a melody.

The pattern of results in both post-hoc analyses indicates an increasing ability to extract pitch information from tones with conflicting spectral information as the level of context increases from isolated to tone series to melodic context. This result was most striking when looking at which F0 differences were detected relative to no F0 change. By this analysis, none of the F0 changes in the isolated condition were discriminated, while all were discriminated in the melodic condition. The tone series context fell halfway between them, as three out of six F0 differences were detected, thus following our third prediction. This result was predicted if the facilitation of pitch discrimination seen in the melodic condition was due to both the presence of other tones and the tonal structure of those tones.

Another difference between the tone-series and melodic conditions is the length of the sequences. The five-tone sequences used in the tone-series condition were shorter in length and number of tones than the melodic sequences. Experiment 3 addresses this potential confound by comparing tone-series contexts differing only in length; a five-tone series and a longer series matching the length of the melodic context. We predicted no difference in the results from these two lengths of tone-series.

EXPERIMENT 3

Methods

Subjects: Five graduate students participated in this experiment. The average length of musical training was 2.6 years. Normal hearing was determined by self-report.

Stimuli: This experiment used two variations in spectral shape, the High and Low timbres described in experiment 1, and two F0 deviations: 0 (same) and 52¢. Stimuli were created and presented through the same computer system as in experiment 1 in a sound-attenuated room.

Procedure: Two variations of the tone-series context described in Experiment 2 were used in this experiment. Context 1 (Short Sequence) used the five-tone alternating and repeating sequences exactly as presented in experiment 2. Context 2 (Long Sequence) extended these series to 34 tones, matching the average length of the melodies in experiments one and two: 8.75 seconds. Each series was presented so that the last note should end on one of six whole steps from C4 to A#4. The last note of each sequence could be played correctly, i.e. at the correct F0, with the same spectral shape at the previous note, or in a different spectral shape and/or F0 than the correct final note. This resulted in four different stimulus types: Same Pitch-Same Timbre, Same Pitch-Different Timbre, Different Pitch-Same Timbre, and Different Pitch-Different Timbre. Subjects were asked to determine if the last note was in tune or out of tune with the rest of the sequence. Eight practice trials were presented before the test conditions, each containing 96 trials (24 trials of each stimulus type).

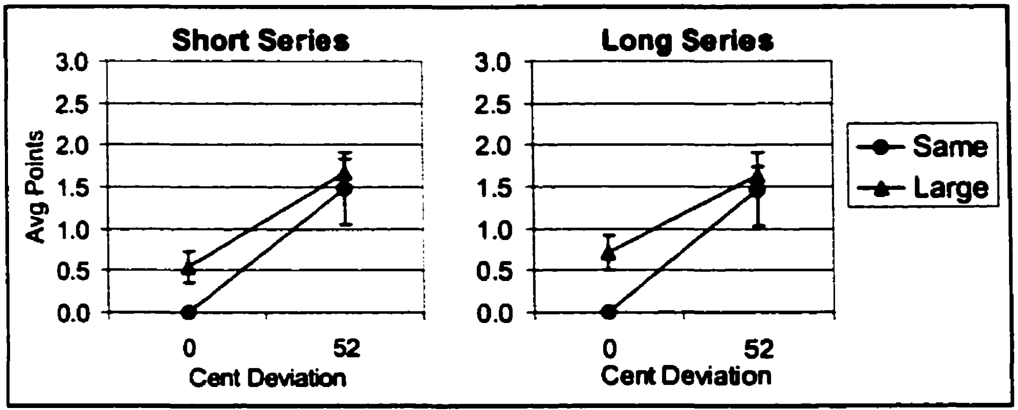

Figure 5: Experiment 3 results averaged across subjects (each subject performed both series). Scored on scale of 0-3 where 0 means "Same Pitch" (See text for details). In timbre differences legend, "Same" denotes same timbre, and "Large" denotes a large difference in timbre. Standard error bars shown.

Results and Discussion – Experiment 3

No differences were seen between responses to the short sequences and the long sequences $F(1,4) = 0.04$, $P>0.8$. No significant interaction between sequence length and stimulus type was found $F(3,12) = 0.55$, $P>0.6$ (See figure 5).

Since no difference was found between responses to short tone-sequences of the length used in experiment 2, and tone-sequences matched in length to the melodies used in experiment 2, the differences seen between the tone-sequence and melodic condition of experiment 2 cannot be attributed to a difference in length.

EXPERIMENT 4

The last experiment in this study addressed the possibility that differences in response instructions could be influencing results. Although the task in each context is to detect deviations in pitch, the semantics of asking this question change between contexts. One responds same vs. different in the isolated context, while responding in tune vs. out of tune in the melodic and tone-series contexts. This experiment addresses this issue and, by using novel melodies, investigates the role that familiarity of the melodies plays in the results seen in experiments 1 and 2. Do the results seen in these previous experiments rely heavily on the fact that familiar melodies were heard, or does the "syntax" of tonality generate a strong enough reference point to obtain the same pattern of results in novel melodies? We predicted that the strict tonality of our novel melodies would create a similar reference point from which to judge pitch as with the familiar melodies in experiments 1 and 2, and that the results from this experiment would follow the same pattern.

Methods

Subjects: Ten graduate students participated in this experiment, five of whom also participated in experiment 3. None had any extensive musical training (average training = 1.6 years). Normal hearing was determined by self-report.

Stimuli: This experiment used two variations in spectral shape, the High and Low timbres described in experiment 1, and two F0 deviations, 0 (same) and 52¢. Stimuli were created and presented through the same computer system as in experiment 1 in a sound-attenuated room.

Procedure: On each trial, subjects were presented with one of four short melodies. When played correctly, the last two notes of each melody had the same pitch (see figure 2 for examples). The last note of each melody could be played correctly, i.e. with the same spectral shape at the same F0 as the previous note, or in a different spectral shape and/or F0, as in experiment 3. The melodies were constructed following strict rules of tonality, and did not present the pitch of the test tones before the end of the melody. When played correctly, each could end on one of six whole steps from C4 to A#4. The participants were divided into two groups, each with slightly different instructions. Group 1 was asked to determine if the last note of each melody was played in tune or out of tune, disregarding any change in timbre. Group 2 was asked to determine if the last two notes of each melody had the same or different pitch, also disregarding any change in timbre. Both groups were presented with identical stimuli, the only difference between groups being the phrasing of the task requirements. Eight practice trials were presented before the 144 test trials, 36 of each stimulus type.

Results

No significant differences were found between the two subject groups or between melodies $F(1,8) = 1.01$, $P>0.3$, and $F(11,99) = 1.31$, $P>0.2$, respectively. Results were

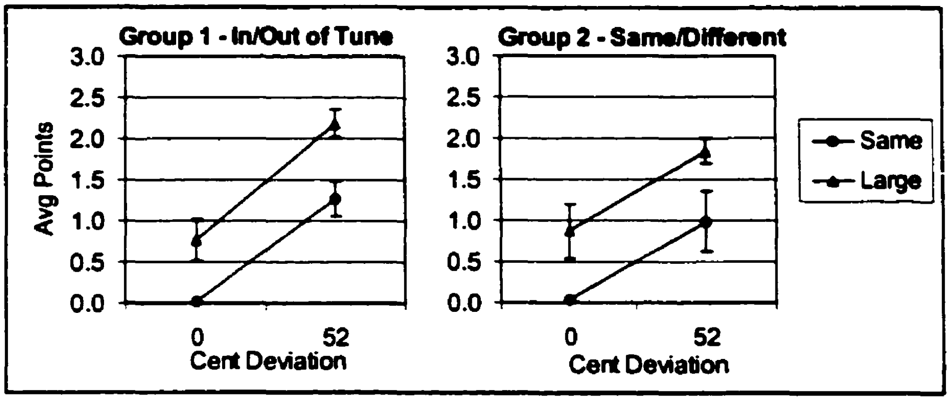

Figure 6: Experiment 4 results averaged across subjects. Each subject participated under one set of instructions only. Scored on scale of 0-3 where 0 means "Same Pitch" (See text for details). In timbre differences legend, "Same" denotes same timbre, and "Large" denotes a large difference in timbre. Standard error bars shown.

consistent with the previous melodic conditions of experiments 1 and 2 (See figure 6). Namely, a main effect of pitch $F(1,9) = 36.04$, $P<0.001$ and a main effect of timbre $F(1,9) = 17.17$, $P<0.005$ were found. Post-hoc examination of which F0 differences were detected, done by comparing the 0 and 52¢ ratings within each timbre pair using Wilcoxon's signed rank test, showed that both 52¢ deviations were detected $Z=-2.803$, $P=0.005$. This pattern is consistent with the melodic conditions of experiments 1 and 2.

Discussion – Experiment 4

No differences were seen between two groups of participants who were asked to judge whether the tones at the end of melodies were played either at the same/different pitch, or in/out of tune. Therefore, the differences seen between conditions of experiments 1 and 2 cannot be attributed to the phrasing of the task instructions. This does not necessarily imply that people are performing the task in the same way in the isolated and melodic conditions. Even though the physical difference upon which each judgment is made is the same in both conditions, there is more information available in the melodic condition. The in/out of tune question is meant to encourage listeners to make better use of the contextual cues that are unavailable in the isolated condition in which an in/out of tune judgment is irrelevant. It is possible that listeners might be able to ignore the context in other cases. However, as seen in this experiment, even when instructed to compare the last two tones of each melody, thus rendering the rest of the melody irrelevant to the task, listeners apparently did not ignore the context as they were able to detect the F0 differences present in the stimuli. This contrasts with results from the isolated conditions in experiments 1 and 2 in which F0 differences were undetectable.

As predicted, the results obtained with novel tonal melodies mimicked the pattern of results seen with familiar melodies in experiments 1 and 2. In those experiments we found the presence of familiar melodies to enhance pitch perception in the face of spectral discrepancies between the test tones. We attributed this result to both the presence and tonal structure of the tones comprising the melodic context. Experiment 4 extends this result to novel melodies, assuring us that it is not merely the familiarity of the melodies that facilitates task performance in the melodic conditions of experiments 1 and 2. The similar results seen with both familiar and novel melodies highlight the robustness of this effect.

General Discussion

This study investigated the interaction of pitch and timbre by examining the extent of the influence of timbre on pitch perception under three different contextual conditions: isolated, tone-series, and melodic. At the small differences in F0 examined in this study, spectral shape always influenced pitch discrimination. In all conditions we examined, participants judged the pitch of different-timbre trials as more different than same-timbre trials regardless of differences in F0. Experiments three and four confirmed that these results were not due to confounding factors such as different lengths of stimuli in the melodic and tone-series conditions, or differences in instruction semantics.

The fact that spectral shape had an influence on pitch judgments indicates that pitch and timbre do interact and are not completely separable dimensions of sound. This result conflicts with the studies using interpolating tones, which find pitch and timbre to be completely separable. However, the interpolated-tone paradigm differs from ours in a couple of crucial aspects. It requires the participant to remember a standard tone over a period of time usually lasting a few seconds, whereas our paradigm allows a more temporally direct comparison. Also, in the interpolated tone studies, test tones were always presented with the same spectral shape. It was the interpolated tones that differed in timbre, and were meant to be ignored by the participants. In our study, the test tones themselves differed in spectral shape; i.e. the test tone could vary in both pitch and timbre from the comparison / correct tone. Ignoring the irrelevant property of timbre in the test tones may be a harder task than ignoring it in irrelevant tones. Indeed, this difference in methodology covaries in the pitch and timbre literature with whether or not test tones are heard in the context of other tones. Perhaps the reason we don't see a complete separation of pitch and timbre in the melodic condition is because we varied both attributes within our test tones.

At larger differences in F0, spectral shape is not likely to affect pitch judgments. Singh and Hirsh's study (1992), which used a paradigm analogous to our isolated condition, pitch and timbre did not interfere with each other at F0 differences of 4% and higher. This 4% difference corresponds approximately to a 70¢ difference, which is larger than our largest F0 deviation of 52¢. Therefore, based on this previous study, we predict that increasing F0 in any of the contexts used in the present study should decrease spectral shape's effect on pitch judgments.

The low level of musical training of our participants could also be affecting the size of spectral shape's effect in this study. A speeded classification study comparing musicians and nonmusicians on pitch and timbre perception showed that nonmusicians' pitch judgments are more affected by timbre differences than musicians' (Pitt, 1994). Trained musicians have more experience not only in making pitch judgments, but determining whether different instruments are in tune with each other, which involves discriminating pitch while ignoring differences in spectral shape. This practice would probably enable musicians to outperform nonmusicians on our paradigm. However, pilot testing in our lab leads us to expect musicians to show the same pattern of results across conditions as the nonmusicians, namely that differences in spectral shape would interfere with their pitch judgments less with increasing tonal context. More research needs to be done to confirm this expectation.

Tonal context appears to be playing an important role in this study. Placing a tone within a tonal context seems to create a better point of reference from which other tones can be judged. Since our Western tonal system is based on ratios, without a reference point listeners cannot determine whether or not a single note is in tune or not (unless they have absolute pitch). A two-tone interval has a specific ratio between F0s, which can be in tune or out of tune. Hearing more notes creates more of a tonal reference point by which one can judge whether each tone is in or out of tune with the rest of the notes present (e.g. Deutsch 1972a; Dewar et al, 1977; Krumhansl 1979). Crowder states that when listening to a sequence of tones, "even when there is no obvious melody, individual tones are not heard independently of a tonal context" (1993, p137). In addition to supporting this model, our results suggest that a tonal reference point can lessen timbre's influence on pitch perception, allowing people to extract pitch information even more efficiently, and that increasing levels of tonal context allow more accurate pitch extraction.

The melodic conditions in this study showed the most facilitating effects, but can we be sure this is due to their tonal structure and not simply their higher structural complexity in comparison to the other contexts? One way to test this idea would be to include a random context, which has a complex structure, arguably more complex than the simple melodies used in this study, and it is atonal. Therefore, if structural complexity was the facilitating factor, task performance should be higher in the random context, and if tonal structure in particular was the facilitating factor, performance should be better in the melodic context. We have made this comparison in another study using a similar task, and found that performance with a random context falls between the isolated and melodic contexts (Warrier et al, 1999). This result argues against the idea that the facilitation we see is due solely to the complexity of the context, and supports the idea that the rich tonal structure of the melodies is facilitating pitch extraction.

Another factor that may be contributing to the differences seen between contexts in this study is the preparatory effect of expectation. Expectations of final notes were created by the tone-series due to their repetitive nature, and by the novel melodies through their strong tonal structure. Familiar melodies, in addition to their tonal structure, created expectancies of the final note simply by being familiar to the subjects; they knew from memory what the final tone should sound like. Expectation causes a mental priming to occur which focuses attention to a particular F0. Bharucha (1994) suggests that this priming involves the formation of a mental image before the last tone is heard, to which the physical final tone can be compared.

One method of testing whether expectancy is playing a role is to test whether a decrease in performance is seen when the stimulus is unexpected. A study looking at the expectancy of chords (Bharucha and Stoekig, 1986) presented chords at the end of a progression, and listeners were asked to decide if that last chord was in tune or out of tune. Performance on unexpected chords was lower than on expected chords. In the present study, the timbre of the last note could be considered expected or unexpected in this way. Hearing a melody or series of tones being played in a specific timbre sets up the expectation that the last note will be played in the same timbre. So the expected timbre in this study was that in which the beginning of the trial was presented (same-timbre trials). When it differed from the beginning timbre, this was unexpected from that trials' context (different-timbre trials). Our listeners were more accurate when the last note was presented in the expected timbre. Evidence such as this suggests the tonal context leading up to the test tones not only creates a tonal framework from which to judge pitch, but also creates expectancies in the mind of the listener which are either fulfilled or violated.

The concepts of melodic context creating a tonal framework, and expectation playing a preparatory role in the melodic condition, both help to explain the differences found between the isolated and melodic conditions in this study, but do not explain the differences found between the two melodies. Although both the "Blue Danube" and "Oh, Susanna" create a tonal framework helpful in discriminating pitch and generate strong expectancies as to what the final note should be, people were slightly more affected by timbre in "Blue Danube" trials. We predicted any difference between the melodies to be in the opposite direction, since "Blue Danube" contains the extra cue of being able to compare the pitch of the final note to those directly preceding it. However, no differences were seen between the repeating and alternating tone series contexts in Experiment 2, indicating that the melodies' endings are not the cause of the differences seen here.

One possible reason for the slightly better results seen in "Oh, Susanna" comes from the smaller interval sizes than those found in "Blue Danube." As Deutsch explains, "there is considerable evidence that melodic sequences are processed more effectively when these are composed of smaller size rather than larger (reflecting the operation of the principle of Proximity)" (1982, p.302; 1978). More evidence is needed to exactly determine the reason a difference between the two melodies was found. However, it is important to remember that although the difference between the two melodies sheds more insight into the nature of the overall interaction between pitch and timbre, this difference is quite small compared to their mutual differences from the isolated condition.

Although we have been describing our results as showing timbre's influence on pitch perception to lessen with increasing tonal context, this could also be interpreted as an increase in the salience of the pitch dimension making it easier to ignore spectral variations. These interpretations are not mutually exclusive. Our paradigm does not allow us to distinguish between the two interpretations, but both imply an interaction between pitch and timbre.

In conclusion, this study demonstrated a progressive improvement in pitch discrimination with three increasing levels of context. The extra tones heard in each trial of the tone series condition, as compared to the isolated condition, created a stronger reference point from which to judge the test tone, thus enhancing pitch discrimination. The facilitation seen in the melodic condition as compared to the tone series condition appears to be due not to the increase in number of tones, but the structured tonality of the melodies. Their rich internal structure lends more cues as to the tonality of the melody than simple repeating or alternating sequences. The auditory system takes advantage of this extra information when extracting pitch both with and without conflicting spectral information.

Chapter 3

The Role of Right Temporal Cortex in the Effect of Melodic Context

One of our most important perceptual functions is the ability to recognize things as being the same under differing conditions, a concept known as perceptual constancy. It has been argued that detecting differences in stimuli is a relatively easy task for the brain to perform, but abstracting similarity from two differing objects is much more difficult (e.g. Whitfield, 1985). Pitch constancy, an auditory example of perceptual constancy, is the ability to judge two tones of differing spectral shape as having the same pitch, even though they do not sound exactly the same. For example, pitch constancy is necessary to hear that a flute and a piano are both playing "middle C," or that two vowels are spoken at the same pitch. To perform this task, it is necessary to extract the fundamental frequency from both sets of harmonics, and compare them.

It is also important to be able to interpret the same event differently under different circumstances. The context in which a sound occurs can change its meaning or importance dramatically. Context can even affect the way we hear and interpret tonal stimuli by affecting pitch perception (Deutsch, 1972a,b; Deutsch & Roll, 1974; Dewar et al., 1977; Krumhansl, 1979; Deutsch, 1982; Krumhansl & Castellano, 1983). Specifically, it appears that placing a tone within the context of a melody enhances the pitch perception of that tone by establishing a reference point from which to judge pitch. Most musical systems, including the Western tonal system, are based on the relationships between tones, particularly their frequency ratios to one another. A two-tone interval has a specific ratio between fundamental frequencies (F0), which can be in tune or out of tune within that system. Hearing more notes creates more of a tonal reference point by which one can judge whether each tone is in or out of tune with the rest of the notes present (e.g. Deutsch, 1972a; Dewar et al., 1977; Krumhansl, 1979). Crowder sums up this idea by stating that when listening to a sequence of tones, "even when there is no obvious melody, individual tones are not heard independently of a tonal context" (1993, p137).

This study explores the neural mechanisms that underlie pitch constancy, as well as how context can affect this perception. We investigated the problem by testing patients with unilateral excisions in the temporal lobe in order to assess the involvement of this area of cortex in these processes, basing our paradigm on a previous study in which we tested neurologically normal listeners (Warrier & Zatorre, Submitted). In that study, we found that in addition to the establishment of a tonal reference point from which to judge subsequent pitches, tonal context can affect pitch constancy, or the extraction of pitch from differing spectral shapes. The results indicated that a difficult pitch constancy task was made easier by placing it in a melodic context. While comparing the pitches of two tones differing in timbre in isolation proved to be very difficult, performance increased dramatically when the same tones were presented within the context of a familiar melody. This result was interesting because it implied that listeners were able to use contextual information to their advantage when performing the pitch extraction necessary for the task.

Spectral analysis plays a large role in pitch constancy. In order to determine that two tones with differing spectral content have the same pitch, it is first necessary to extract the fundamental frequency from the spectral content of each tone. Another task that involves comparing the spectral composition of tones is timbre discrimination. In this case, the task is not to extract underlying information (such as pitch) from the tones, but instead to judge them on their overall sound quality. This task has been shown to be problematic for patients with right temporal lobe lesions (Milner, 1962; Samson & Zatorre, 1994). Since the auditory cortices are located along the superior temporal gyrus, this suggests that the neural computations necessary for spectral analysis are performed in right auditory cortex.

Lesion studies examining the neural correlates of pitch processing also find patients with right temporal lobe lesions to have difficulty (Milner, 1962; Samson & Zatorre, 1988; Sidtis & Volpe, 1988; Zatorre, 1988; Divenyi & Robinson, 1989; Peretz, 1990; Zatorre & Samson, 1991; Liégeois-Chauvel et al., 1998; Johnsrude et al., 2000). Simple frequency discrimination tasks can be performed by patients with either right or left temporal lobe lesions (e.g. Milner, 1962; Zatorre, 1988; Johnsrude et al., 2000). Even bilateral lesions of auditory cortex don't cause permanent impairments on this task (Peretz et al., 1994). Simple discriminations of frequency (but not necessarily pitch) can be performed without the auditory cortex at all (See Whitfield, 1985, for review). So it appears that comparing F0 when spectral shape stays constant is not affected by lesions in auditory cortex. However, patients with right temporal lobe lesions do show impairments on other types of pitch processing tasks.

For example, in a tonal memory study, patients with right or left temporal lobe lesions performed as well as controls in a simple pitch discrimination task (Zatorre & Samson, 1991). However, when interfering tones were inserted between the standard and comparison tones, patients with right but not left temporal lobe lesions were impaired compared to controls. Discriminating melodic pitch patterns also appears to cause patients with right temporal lobe lesions difficulty (Milner, 1962; Samson & Zatorre, 1988; Peretz, 1990; Liégeois-Chauvel et al., 1998).

In another study investigating the role of temporal cortex in tone perception, patients with right or left temporal lobe lesions performed at the same level as controls on a frequency discrimination threshold test (Johnsrude et al., 2000). However, when asked to determine the direction of the pitch change in the same stimuli, patients whose right temporal lobe lesions included Heschl's gyrus (primary auditory cortex) were impaired. Similarly, no difficulties were observed in another simple frequency discrimination task using complex tones, but when the harmonic corresponding to the fundamental frequency was removed from each tone, patients with right temporal lobe lesions that included Heschl's gyrus were impaired (Zatorre, 1988). This study implicated right Heschl's gyrus in extracting pitch from tones widely differing in spectral shape.

Functional neuroimaging studies conducted on normal listeners support the lesion data in that tasks involving pitch processing usually show right temporal and frontal activations (Zatorre et al., 1992; Zatorre et al., 1994; Halpern and Zatorre, 1999; Perry et al., 1999; Warrier et al., 1999). The frontal activations are generally understood to show the upkeep of tones in memory while performing the task. Zatorre and Samson's tonal memory study (1991) corroborates this view in that patients whose lesion included right frontal areas also showed impaired performance on the pitch discrimination task when interference tones were interpolated between the standard and comparison tones.

The present study investigates the neural correlates of pitch constancy and context effects by testing patients who have had unilateral temporal lobe excisions. The goals of the study were two-fold. First, we wanted to investigate the role of right auditory cortex in the spectral processing of tones. For this we tested patients with temporal lobe excisions and a control group on a pitch constancy task. Based on previous studies showing that patients with right temporal lobe lesions are impaired on many pitch-related tasks involving more than the most simple of frequency comparisons (see above), we predicted that they would also have difficulty with the pitch constancy task. More specifically, due to Zatorre's 1988 study using tones with missing fundamentals, we predicted that patients whose right-sided lesions included Heschl's gyrus would be impaired when the task required extraction of pitch from tones of differing spectral shape, as when determining pitch constancy, but not when a direct comparison was available and this process was unnecessary.

Secondly, if patients with right temporal lobe lesions exhibited the predicted deficit in spectral processing, we were interested in examining whether or not they would be able to utilize tonal information contained in contextual tones to facilitate this processing. Therefore, we used the paradigm of Warrier and Zatorre (submitted) that showed melodic context to facilitate pitch extraction from tones differing in spectral shape in neurologically normal subjects. We predicted that the contextual information would not aid the patients with right temporal lobe lesions on this task.

Methods

Participants: All thirty-six patients tested had undergone focal unilateral cerebral excision at the Montreal Neurological Hospital for relief of pharmacologically intractable epilepsy, with 18 temporal lobe excisions on the left (LT), and 18 on the right (RT). Patients were excluded from the study if they presented atypical speech representation, known damage outside the region of surgical excision, EEG abnormality contralateral to the side of the lesion, a malignant tumor, Full-Scale WAIS-R IQ (Wechsler Adult Intelligence Scale – Revised) of under 75, or evidence of hearing loss or impairment. Twelve neurologically normal control participants (NC), matched to the patients with respect to age and level of education, were also tested (See Table 1 for demographic details). Musical experience was generally limited in each group, with 79% of all participants having two years or less training and/or experience. Although the LT group contained slightly more people with musical backgrounds (four people with more than 8 years experience), a statistical comparison showed no difference between groups on this measure. Testing was performed in each participant's preferred language, either French or English. The ethics committee of the Montreal Neurological Institute approved the experimental protocol, and written informed consent was obtained from all participants before testing.

All resections included the amygdala, uncus, and anterior temporal lobe in one hemisphere. The extent of resection along the hippocampus and parahippocampal gyrus varied from patient to patient, as did the extent of the lateral neocortical excision along the sylvian fissure, the second temporal gyrus and the base of the temporal lobe. Patients were divided according to side and whether or not the lesion extended posteriorly into Heschl's gyrus, a posterior region on the superior temporal gyrus that contains the primary auditory cortex.

Lesion classification was performed according to a method described in Penhune et al (1999). First, each patient's post-operative MRI (magnetic resonance imaging) scan was transformed into standardized stereotaxic space (Talairach & Tournoux, 1988). These scans were then co-registered with a probabilistic map of the primary auditory cortical region developed by Penhune and colleagues (1996), which allows identification and estimation of removal in this area. Of the 18 patients in the RT group, 7 had lesions encroaching upon Heschl's gyrus. The remaining 11 patients' lesions stopped anterior to this area (see Figure 1). Of the 18 patients in the LT group, 3 had lesions encroaching upon Heschl's gyrus, while the remaining 15 had lesions that stopped anterior to this area (see Figure 1).

	N (M/F)	Age	Years of Education	Years of Music Experience	Full Scale IQ (WAIS-R)
NC	12 (3/9)	36.8 (22-52)	13.2 (9-17)	1.2 (0-5)	-
LT	18 (7/11)	35.4 (22-49)	13.5 (10-18)	4.5 (0-15)	106.4 (89-135)
RT	18 (14/4)	36.9 (24-53)	12.8 (8-19)	1.9 (0-15)	101.6 (82-120)

Table 1: Ranges are indicated in parenthesis, except in first column where they indicate gender.

Figure 1: Reprinted from Penhune et al (1999). The upper panel illustrates the MRI scan of a patient with a removal in the right primary auditory cortex region (Heschl's gyrus) where the excision includes the anterior-lateral 50-60% and the undercutting extends to 60-70%. The scan is presented in the horizontal (A: $z = 4$) and coronal (B: $y = -17$) planes of section. The images on the left show the patient's scan alone with an arrow indicating the region of excision/undercutting. The images on the right show the patient's scan coregistered with a probabilistic map of Heschl's gyrus (scaled from 25-100% probability). The cursor indicates the same position in standardized space as the arrow. The box in the lower left corner indicates the region of the removal pictured in close-up in the lower panel. The lower panel shows four close-up views of the same scan, illustrating the transition from intact, to undercut, to fully excised tissue (coronal sections taken at 3mm intervals). Arrows indicate the location of the Heschl's gyrus region. In the section labeled 'undercut' note the very thin band of white matter connecting the remaining portion of Heschl's gyrus to the rest of the brain. In the section labeled 'excised & undercut' note the unexcised strip of presumed gray matter tissue.

Stimuli: The stimuli were digitally synthesized on an IBM compatible 486 computer using MITSYN software (Henke, 1981). All test tones were 500 msec duration, with rise and fall times of 10 msec. Contextual tones used in the melodic condition varied in length according to the necessary tone durations. Three timbres were created by varying the relative intensities of 11 harmonics. One emphasized the lower harmonics (Low), another emphasized the middle harmonics (Middle) and sounded brighter than Low, and the last emphasized the higher harmonics (High) and sounded brighter than Middle (see Figure 2). Individual tones ranged in fundamental frequency from 164.81 – 1108.70 Hz. (E3 to C#6), with test tones always presented between 261.63 Hz (C4) and 480.35 Hz (A#4 + 52¢). Stimuli were presented with MAPLE (McGill Auditory Perception and Linguistic Experiments) software (Achim et al., 1992). Sounds were presented binaurally at 75 dB SPL over Sennheisser HD424 headphones.

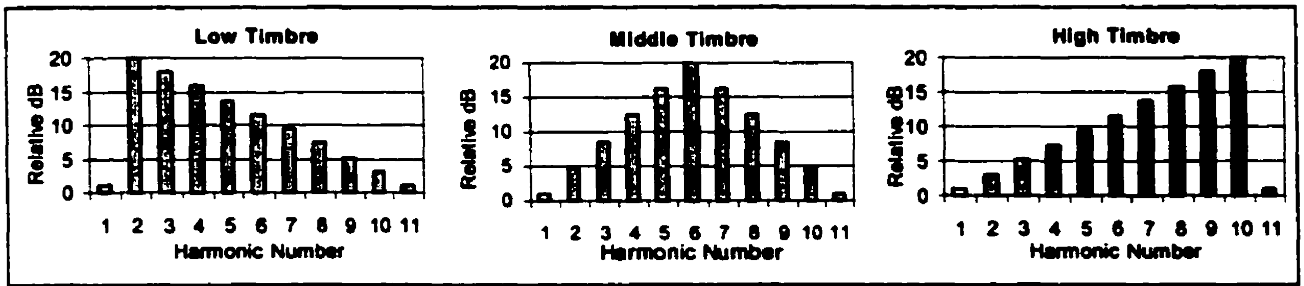

Figure 2: Spectral shapes of timbres. Differences in intensity level of each harmonic are shown relative to the other harmonics.

Procedure: The testing session was broken down into three sections: 1) Explanatory phase, 2) Isolated condition, and 3) Melodic condition, in that order. The entire session, conducted in a sound-attenuated room, took between 60 and 75 minutes.

Explanatory Phase: In order to ensure that everyone understood the concepts relevant for the task, an explanatory phase was conducted with a keyboard before the testing. Pitch was explained by playing different notes on the keyboard. An informal preliminary quiz was then performed, with the participant responding whether two notes had the same pitch or not. The quiz began with large intervals (7th) and gradually worked down towards smaller intervals (2nd). The concept of timbre was then explained as "instrument sound" by playing two notes of the same frequency at different instrument sounds available on the keyboard. After presenting a variety of exemplars, a second quiz was given in which two tones of different instrument sounds were played, and the participant determined if the pitch stayed the same or not, disregarding the change in instrument. Interval size was again varied from large to small. This informal testing was continued until the experimenter was confident that the participant understood the concepts of pitch and timbre.

Isolated Context: In the isolated context condition, each trial consisted of two tones presented with an interstimulus interval (ISI) of 100 msec. The first tone was randomly presented at one of six frequencies corresponding to whole-tone steps starting at C4 (261.63 Hz) and continuing to A#4 (466.16 Hz). The second tone was either 0 (same), 35, or 52¢ higher than the first tone. Cents are logarithmically equal steps in the frequency dimension, in which equal cent differences are perceived as roughly equal pitch differences in any frequency range. Each semitone of the musical scale is 100¢, while 50¢ corresponds to a quarter-tone. Each pair of test tones could be presented in one of three timbre pairings: 1) same (Low-Low, High-High), 2) small difference (Low-Middle, High-Middle), and 3) large difference (High-Low, Low-High). Each pairing was presented at all levels of F0 difference. Participants were informed that they would be making pitch judgments on these tones, and that they should ignore any timbre changes as much as possible. On each trial, participants were instructed to indicate whether the pitch of the second tone was the same as or different from the first tone by pressing a key on the computer keyboard. Subjects completed 54 trials at their own pace after
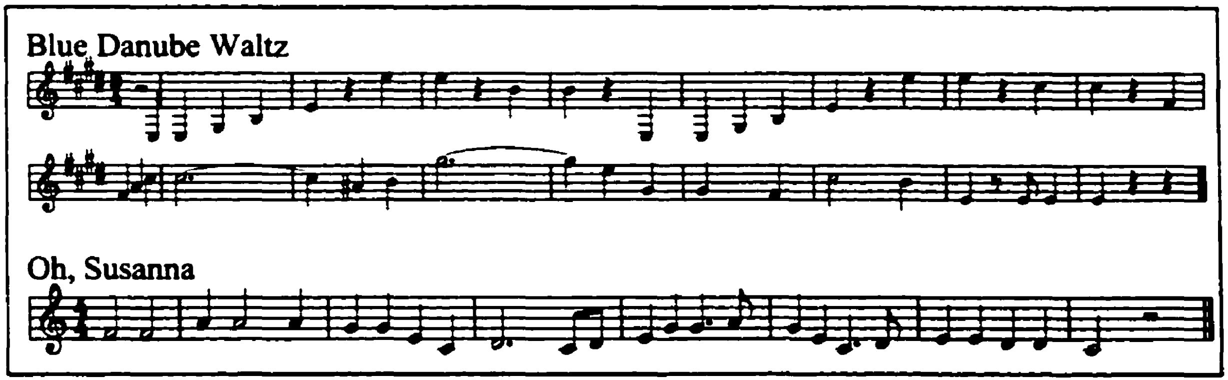
Figure 3: Melodic contexts shown in musical notation.

completing a block of 12 practice trials.

Melodic context: In the melodic context condition, the task was to determine whether the last note of a melody was in tune or not. One of two melodies, familiar to all participants, was presented on each trial: "Oh, Susanna," or "The Blue Danube Waltz" (See figure 3). Each melody, excepting the last note, was presented in tune in one of three keys so that the last note should either end on one of six whole steps from C4 to A#4, and was presented in one of the three timbres described above. The last note could continue with the same spectral shape as the melody, or sound in one of the other two, thus creating the same three timbre pairs as the isolated condition (same, small difference, and large difference), and was 0 (same), 35, or 52¢ sharper than the correct final note. After each melody, participants were instructed to indicate if the last note of the melody was in or out of tune, ignoring any timbre differences. Subjects completed 54 trials of each melody at their own pace, after a block of 12 practice trials.

Results

We calculated percent correct scores from the data and performed all stats on this measure. d' calculations are not reported here. Due to the time constraints of the testing sessions, we were not able to collect enough replications within each cell from each participant to make the d' data meaningful.

The NC group replicated our previous study in that the isolated condition was performed near chance, while scores improved substantially in the melodic condition (Warrier & Zatorre, Submitted). A three-factor analysis of variance (ANOVA) was

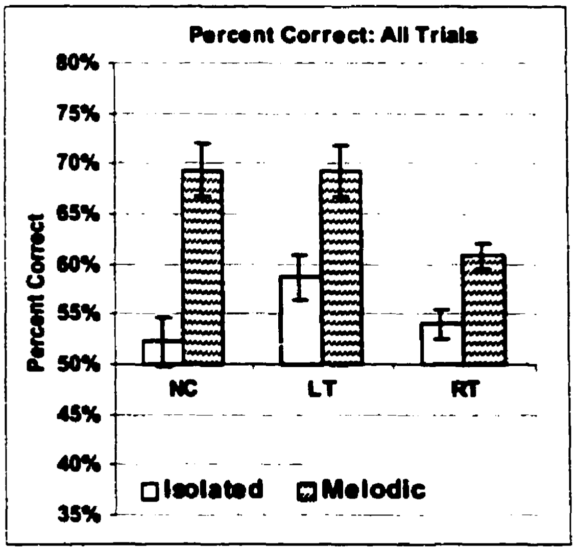

Figure 4: Percent correct results by condition averaged across same- and different-timbre trials. Error bars are indicated.

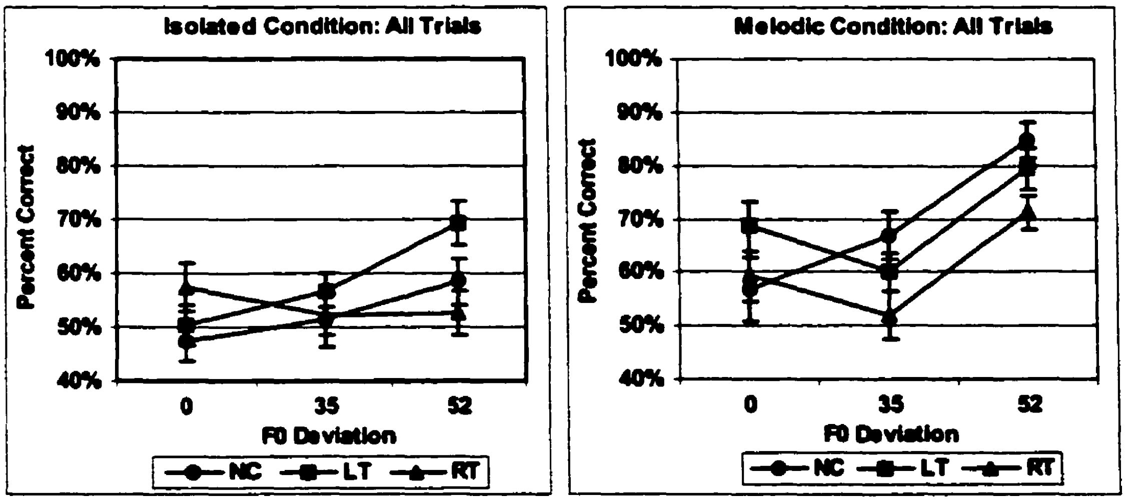

Figure 5: Percent correct results by pitch averaged across same- and different-timbre trials. Isolated and melodic conditions are shown separately. Error bars are indicated.

performed with two levels of Context (isolated and melodic), three levels of Pitch (0, 35, and 52¢), and three levels of Group as a between subjects factor (NC, LT, and RT). Three main effects were found. A main effect of Context $F(1,45) = 83.03$, $P<0.001$ showed that scores were higher overall in the melodic condition (See Figure 4), and a main effect of Pitch $F(2,90) = 10.68$, $P<0.001$ showed that scores generally increased with increasing F0 deviations (See Figure 5). A main effect of group was also found $F(2,45) = 3.74$, $P<0.05$, along with three interaction effects: a two-way interaction between Context and Pitch $F(2,90) = 5.99$, $P<0.005$, a two-way interaction between Context and Group $F(2,45) = 5.10$, $P=0.01$, and a three-way interaction between Context, Pitch, and Group $F(4,90) = 4.24$, $P<0.005$.

We then broke down the three-way interaction by testing for group differences at each level of pitch difference, looking at each condition separately with Tukeys HSD post-hoc tests. The LT group did not differ from controls at any F0 deviation in either condition. The RT group did not differ from controls at any F0 deviation in the isolated condition, but scored significantly lower than controls on both the 35 and 52¢ deviations in the melodic condition, ($P<0.05$ for both comparisons).

We tested whether or not the groups differed in degree of improvement from the isolated to melodic condition with a one-way ANOVA, and found a significant effect of group $F(2,45) = 5.10$, $P=0.01$. Post-hoc Tukey tests revealed no difference in improvement between the LT and NC groups, and a decreased level of improvement compared to NC in the RT group, $P<0.01$. Despite this result, all groups showed a significant improvement from the isolated to melodic context $t(11) = 6.64$, $P<0.001$; $t(17) = 4.49$, $P<0.001$, $t(17) = 4.41$, $P<0.001$ for NC, LT, and RT respectively, computed as Bonferroni-corrected t-tests.

In order to examine the contribution of Heschl's gyrus in this study, the RT group was split into two smaller groups; 7 patients whose excision extended into Heschl's gyrus, and 11 in whom this gyrus was spared. A three-factor analysis of variance was performed with two levels of Context (isolated and melodic), three levels of Pitch (0, 35, and 52¢), and two levels of Group as a between subjects factor. No differences were found between these two subgroups.

Restricting the data to include only same-timbre trials enabled us to remove the spectral analysis factor of the study. These trials required a simple comparison of F0 between two tones that did not differ in spectral shape. Looking at these trials allowed us to focus on whether the patients with right temporal lobe lesions were able to do the task at all, and if they were utilizing the contextual cues found in the melodic condition. Because timbre influenced judgments so much at the small differences in F0 we used in this study, same-timbre trials were biased to be judged as same in pitch. This accounts for the low percent correct scores seen in all groups on the 35 and 52¢ trials in the same-timbre trials in the isolated condition, while 0¢ deviation scores were near 100% correct (See Figure 7).

A three-factor analysis of variance was performed on the same-timbre trials with two levels of Context (isolated and melodic), three levels of Pitch (0, 35, and 52¢), and three levels of Group as a between subjects factor (NC, LT, and RT). Three main effects were found. The main effect of context indicated higher performance in the melodic condition $F(1,45) = 72.18$, $P<0.001$ (See Figure 6), a main effect of pitch was found to verify the increase in scores with higher F0 deviations $F(2,90) = 268.49$, $P<0.001$ (See Figure 7). A main effect of group was also found $F(2,45) = 7.98$, $P=0.001$. Two interaction effects were found: an interaction between Pitch and Group $F(4,90) = 4.35$, $P<0.005$, and between Context and Pitch $F(2,90) = 38.03$, $P<0.001$.

We again broke down the interaction effects by testing for group differences at each level of pitch difference within each condition separately using Tukeys HSD post-hoc tests. In the isolated condition, the only difference from controls was found in the LT group on the 52¢ trials, in which the LT group scored higher than the controls $P<0.05$. In the melodic condition, no differences were found between LT and NC groups. However, in the melodic condition, the RT group scored significantly lower than controls on both the 35 and 52¢ conditions, $P<0.05$ for both comparisons.

Within the RT group, we again separated those patients whose iesions encroached upon Heschl's gyrus and those in whom this area was spared, and looked at same-timbre trials only. A three-factor analysis of variance was performed on these two groups, with two levels of Context (isolated and melodic), three levels of Pitch (0, 35, and 52¢), and two levels of Group as a between subjects factor. No differences were found between

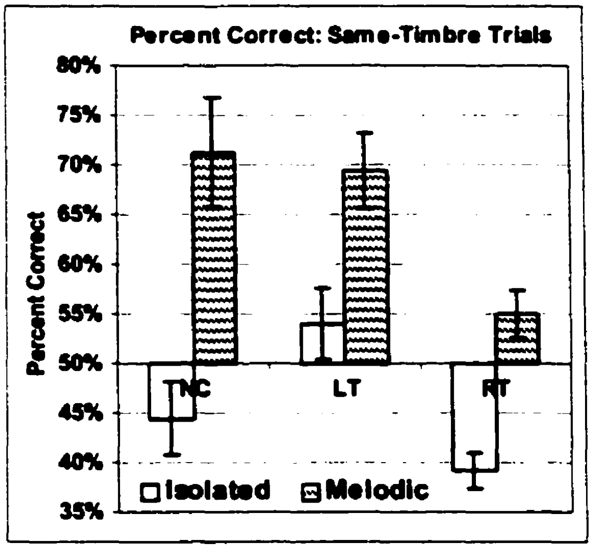

Figure 6: Percent correct results by group and condition on same-timbre trials only. Error bars are indicated.

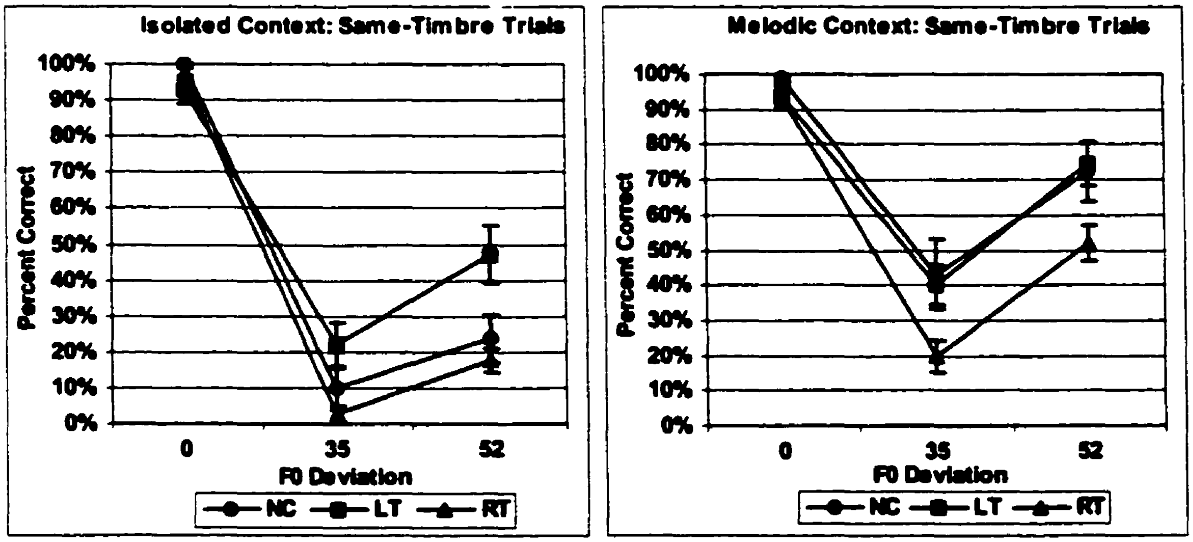

Figure 7: Percent correct results by pitch on same-timbre trials only. Isolated and melodic conditions are shown separately. Error bars are indicated.

these two groups. When comparing the same two groups in the same manner with different-timbre trials only, again no differences were found.

Discussion and Conclusions

As in our previous behavioral study, all subjects had difficulty performing the pitch extraction task in the isolated condition. Placing the test tones within the context of a melody improved performance for all three groups, but the degree of improvement varied between groups. In particular, patients with right temporal lobe damage did not improve as much as controls, while patients with corresponding lesions in the left hemisphere achieved the same level of improvement as controls. Within our RT group, performance did not differ if the lesion included or excluded Heschl's gyrus.

Our RT group did not show the expected effect of spectral shape. The trials which differed in timbre required extraction of pitch from differing spectral shapes. We predicted that the RT group would have an easier time with the task in same-timbre trials when this extraction was not necessary. To assess effect of spectral shape directly, we looked at scores in the isolated condition only. If spectral processing was a factor in this study, we should have seen better performance by the RT group on same-timbre trials in the isolated condition than when different-timbre scores were included in the analysis. In fact, the RT group's scores did not differ from the NC group's in either analysis. In addition, contrary to our prediction, no effect was seen on same- or different-timbre trials when dividing the RT group into two subgroups: patients whose excisions extended into Heschl's gyrus or not. However, these non-results may be due to a floor effect seen across groups in the isolated condition in which all groups, including NC, performed only at around 50% correct, or chance. So although no effect of spectral shape was seen in this study, we cannot conclude that the RT group's, or subgroup's ability to process spectral shape is in no way impaired.

The difficulty shown by all groups in the isolated condition mimics the data from a previous behavioral study in which normal listeners were unable to ignore timbre differences when judging differences in pitch in that condition (Warrier & Zatorre, Submitted). At the small frequency differences used in both this and the previous study, differences in timbre were interpreted as differences in pitch such that same-timbre trials were judged as having the same pitch, and different-timbre trials were judged as having different pitch, whether the pitch changed in that trial or not. It is only in the melodic condition that, although pitch judgments are still influenced by timbre, listeners are better able to focus on the frequency dimension, and extract pitch from the spectral information of the tones. The present study showed that the capability of melodic context to help listeners ignore timbral differences is considerably weaker in the RT group. Therefore, this group was not able to use the facilitating information contained in the melodic context to the same extent as controls or patients with left temporal lesions. In order to make full use of the melodic context, it is necessary to consider pitch relative to contextual tones that were heard previously. Since our RT group had difficulty on this task while the LT group did not, it appears that at least some of the processing necessary to fully utilize tonal context in this way is done in the right temporal lobe. It is in this area of cortex that we propose tones are analyzed with respect to those that came previously – a process that must of necessity involve a tonal working memory aspect.

A tonal memory impairment may be contributing to the decreased scores seen in the RT group. Previous lesion literature has shown patients with right temporal lobe lesions to exhibit impairments when the task requires holding tones in memory over a short period of time (Zatorre & Samson, 1991; see also Chao and Knight, 1997). Therefore, our RT group may be having difficulty holding the contextual tones in memory long enough for them to affect perception of the test tones, and are thus not obtaining the full benefit of the context.

The RT group's impairment may also be due in part to an inadequate processing of the melodies. By not being able to fully process the melodies that make up the melodic context, these patients would not have all the melodic information available to them that the LT and NC groups did. Previous research has shown patients with right-hemisphere lesions to exhibit impaired performance on tasks requiring them to process melodic contours or intervallic information within the melodies (Zatorre, 1985; Peretz, 1990; Liégeois-Chauvel et al., 1998). However, these studies also find patients with left hemisphere lesions to have difficulty processing intervallic information. Therefore, difficulty processing the melodic contour, a deficit apparently specific to patients with right-hemisphere lesions, is a more likely candidate for contributing to the decreased performance in the RT group. By not processing contour information effectively, the RT group may not obtain the same facilitating information from the contextual melodies as the other groups did.

One of the main goals of this study was to assess the different levels of involvement of primary auditory cortex vs. the more anterior auditory areas in spectral analysis and utilizing tonal contextual cues. As discussed above, no dissociation of responses was found between RT patients whose excision included or spared Heschl's gyrus, which contains the primary auditory cortex. The deficits seen in the RT group can therefore be attributed to the excision of more anterior areas of the temporal lobe, excised in all RT patients. This implies that anterior portions of the auditory cortex are involved in making pitch judgments relative to tones that were heard previously.

The auditory association areas AA_2 and AA_3 are located on this anterior portion of the superior temporal gyrus, with AA_3 located anterior to AA_2 (Chavis & Pandya, 1976). These areas have cortico-cortical connections to the superior temporal sulcus, as well as frontal and paralimbic areas (summarized by Pandya & Yeterian, 1985; Pandya, 1995). Area AA_2 is particularly interesting in relation to the present study due to its connections to areas in the frontal lobe known to be involved in working memory (Petrides & Pandya, 1999; Romanski et al., 1999a). These connections could be involved in holding the contextual tones of melodic context in memory while listeners wait for the final tone.

Supporting this view is a tonal memory study that tested patients with anterior temporal and/or frontal lesions (Zatorre & Samson, 1991). The task involved comparing the pitches of the first and last notes of a series of tones, requiring that the first tone be held in memory while the sequence played out. Patients with right temporal or frontal lesions were impaired on this task. A recently conducted fMRI study compared blood oxygenation level dependent (BOLD) responses when listeners made same/different pitch judgments in different tonal contexts also supports this idea (Warrier et al., 1999a). When subtracting BOLD response to judgments made in an isolated context from those made in a melodic context, contexts similar to those used in the present study, only one area of cortex showed a significant difference: the anterior portion of the right superior temporal gyrus, an area excised in all of our RT patients. Given this convergence of results, we feel more confident in stating that the neural computations necessary for bringing about the facilitating effects of tonal context are performed in this anterior area of the auditory cortex.

In conclusion, this study examined perceptual constancy from an auditory perspective. We used pitch constancy as an example of perceptual constancy, and explored how context can affect this perception. In particular, we investigated how the auditory cortex can use contextual cues to facilitate pitch constancy. Although a melodic context greatly facilitated pitch constancy in both controls and patients with left temporal lobe lesions, it helped patients with right temporal lobe lesions much less. We propose that the auditory processing necessary for judging tones relative to those that were heard previously involves neurons located in right anterior auditory cortex. This area appears to work in conjunction with right dorsolateral frontal cortex, an area previously implicated in tonal working memory.

Chapter 4

Right Anterior Superior Temporal Gyrus Implicated in Tonal Context Effects

An important feature of our perceptual system is that it can interpret the same event differently under different circumstances. For example, the context surrounding an event can change its meaning or importance dramatically. Contextual effects can be produced by co-occurring stimuli as well as events separated in time. The processes generating these effects are manifested in behavioral differences to the same stimulus when the surrounding stimuli change. One hypothesis is that the perceptual system is able to produce these contextual effects through a hierarchy of stimulus processing. Lower levels of this system extract stimulus features, such as frequency, edges, or scent, while higher levels are able to take this information and perform additional processing, such as relating similar aspects of stimuli to each other.

One example of contextual effects occurs in the auditory modality: placing tones within a tonal context has been shown to affect pitch discrimination, implying that the neural processing of tones can be influenced by tonal information heard previously (Deutsch, 1972b; Dewar et al., 1977; Krumhansl, 1979; Krumhansl & Castellano, 1983). Specifically, behavioral data show that listeners discriminate pitch more accurately when tones form part of a meaningful structured tonal pattern (a melody) than when they are presented in isolation (Warrier & Zatorre, Submitted).

The present study uses fMRI technology to investigate where within the cortical auditory processing regions the effect of tonal context occurs. The auditory system is organized in a hierarchical manner with core (primary) areas surrounded by belt areas (Rauschecker et al., 1995). We hypothesized that the contextual effect would require the participation of a region located in these neighboring association areas. A previous behavioral study showed that patients with anterior removals of right superior temporal gyrus (STG) were impaired in the use of melodic context to facilitate a pitch discrimination task when compared to a neurologically intact control group (Warrier & Zatorre, 1999). This suggests that anterior auditory regions in the temporal lobe are able to process tones in light of previously presented tonal information. In order to accomplish this processing, one would expect anatomical connections to a working memory system that could hold previously presented tones in memory while processing incoming tones. The dorsolateral frontal cortex has been implicated in many different types of working memory tasks (Owen, 1997; Petrides, 1998), and involvement of right frontal cortex specifically has been demonstrated in tonal working memory tasks with both lesion and functional imaging studies (Zatorre & Samson, 1991; Zatorre et al., 1994; Perry et al., 1999). Anterior auditory association areas on STG have cortico-cortical connections to dorsolateral frontal cortex (Romanski et al., 1999a), making the right anterior STG a good candidate for conveying auditory contextual effects.

In the present study, we scanned 10 musicians while they were making pitch discriminations under three different tonal contexts: 1) isolated, 2) random, and 3) melodic. Each condition was compared to silence as well as the other conditions. We compared low context and high context conditions to isolate areas involved in producing the context effect, predicting increasing activity in the right anterior STG and right dorsolateral frontal regions as a function of increasing context.

Methods

Subjects: Ten right-handed musicians aged 19 to 31, with 0-12 years of formal musical training (average 5.7 years), and 5-21 years of practice (average 13.9 years) participated in this study. Written informed consent was obtained from each subject, and all were compensated monetarily for their participation. One scanning session produced an unsatisfactory amount of artifact, causing that person's data to be unusable. In addition, one control scan file was truncated during transfer onto disk, causing that scan to be unusable. We therefore included data from 9 people (6 male, 3 female) in the test analyses, and 8 in the control condition.

Stimuli: Stimuli were digitally synthesized on an IBM compatible 486 computer using MITSYN software (Henke, 1981). All test tones were 400 msec duration, with rise and fall times of 10 msec. Contextual tones varied in duration from 250 to 650 msec with rise and fall times of 10 msec. The waveform of all tones was sawtooth with 15 harmonics. Fundamental frequency of the tones ranged from 196.00 Hz (G3) to 622.25 Hz (D#5). The first test tone was always presented at either C4, D4, or E4 (261.63, 293.66, or 329.63 Hz). The second test tone was presented either at the same frequency

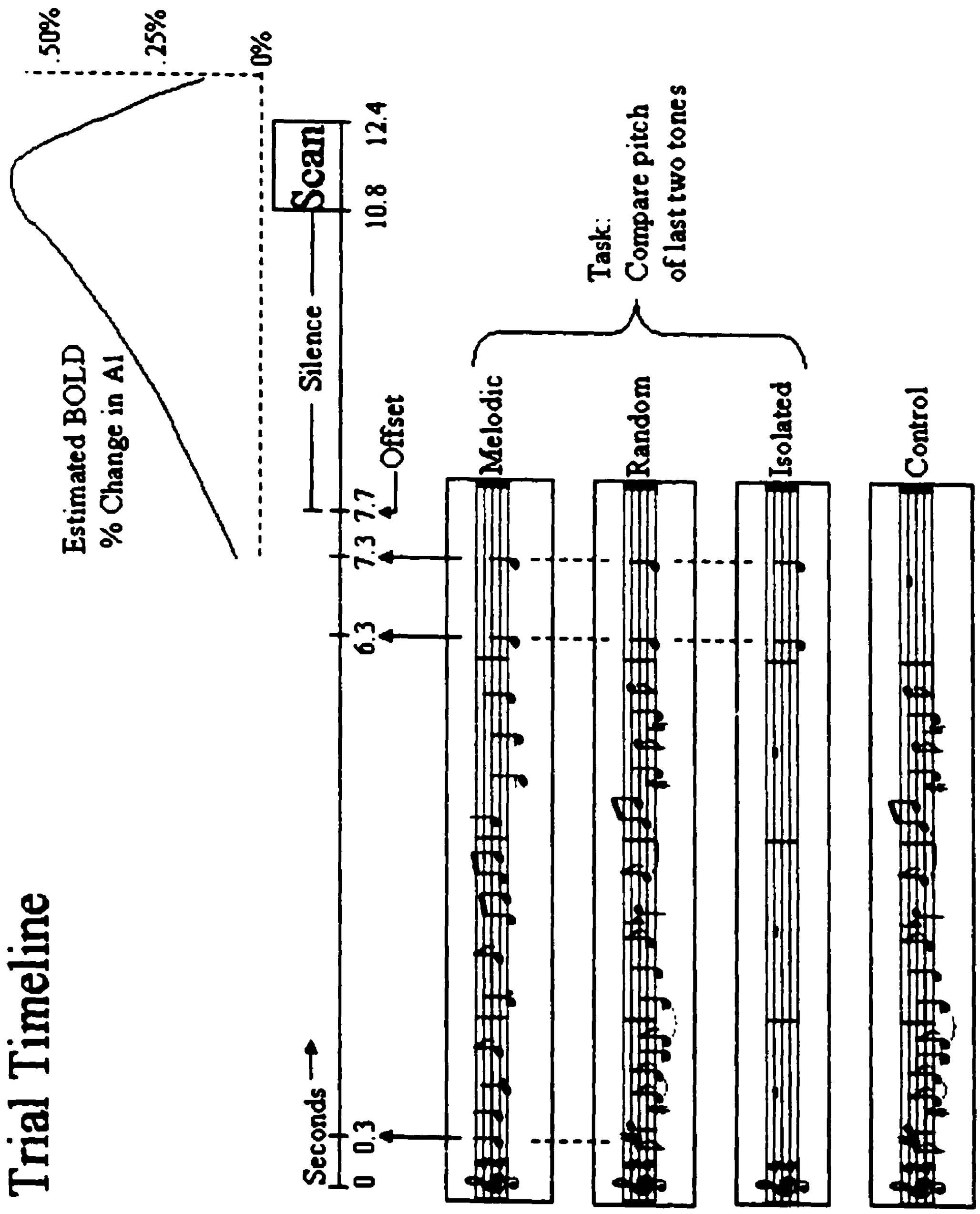

Figure 1: Schematic representation of timecourse for each condition. Time is shown in seconds increasing from left to right. Tonal stimuli are shown in musical notation for ease of demonstration. Presentation of contextual stimuli began 0.3 seconds after end of previous scan, and sounded for 6 seconds. Test tones of 400 msec duration were presented at 6.3 and 7.3 seconds from the beginning of the trial. All 16 scan slices were obtained within the 1.6 second interval beginning 10.8 seconds from the beginning of the trial. A chart of the timecourse of estimated percent change in BOLD signal due to the second test tone is shown above the trial timeline (adapted from Belin et al., 1999).

(same-pitch trial), or plus or minus 1/5 of a semitone more than the individual's pitch discrimination threshold as determined by a previous testing session (different-pitch trial). Stimuli were presented with MAPLE software (Achim et al., 1992) at 80dB SPL through Koss E/90 electrostatically shielded headphones.

Behavioral task: Three pitch discrimination tasks were presented, each differing only in tonal context: Compare the pitches of two tones preceded by 1) silence (Isolated), 2) a random series of tones (Random), or 3) a melody (Melodic) (See Figure 1). The random and melodic contexts each sounded for 6 seconds before the onset of the first test tone, and were matched in terms of individual tone durations and frequency range. Within each trial, no contextual tone was presented at the frequency of the test tones. Melodies were constructed in order to follow these constraints and to lead tonally to the test tones as much as possible within the western system of tonality. Random tone series were computer generated by pseudo-randomly choosing from lists of tones, then presenting them in a random order. Thirty-two trials of each condition were randomized along with 31 silence trials, and presented in the same random order for each subject. Listeners pressed a mouse button with their right hand after each trial indicating whether the two test tones were same or different in pitch.

In a separate session during the week prior to scanning, we determined each subject's pitch discrimination threshold with a behavioral test using a staircase method(Levitt, 1971). The different-pitch task stimuli were then set to 1/5 of a semitone above individual thresholds in order to control for task difficulty (range: 25-55☛, average: 43.5☛; semitone=100☛). During this session the task was described and practice trials were administered.

Functional methods and statistical analyses: fMRI images were obtained with a 1.5 Tesla Siemens Vision scanner (head coil), using Echo Planar Imaging with clustered image acquisition. Sixteen 7mm thick slices oriented parallel to the anterior/posterior commissure line were obtained every 12.4 seconds, with matrix size 64 x 64, and voxel size 5 x 5 x 7 mm^3. These slices were obtained in rapid succession, 100msec per slice, totaling 1.6 seconds per scan (See Figure 1). Anatomical scans were obtained from each subject for co-registration with the functional data. Individual scans were motion corrected, smoothed with a 6mm FWHM filter, and resampled into standardized

Talairach space (Talairach & Tournoux, 1988) for group analysis. The data were analyzed with Worsley's fMRIstat software based on a linear model with correlated errors (Worsley, 1999). Activation peaks were considered significant at or above $t = 4.82$ ($p < 0.05$) for whole brain search (Worsley et al., 1996). The critical t was lower for more directed searches, depending on the size of the region of interest.

Coordination with the scanner: Since we were investigating differences in auditory processing between test tones presented in differing contexts, we wanted to measure the blood oxygenation level dependent (BOLD) response due to processing of the test tones, and not due to listening to the contextual tones. We therefore used a method developed by Belin and colleagues (Belin et al., 1999) that uses information regarding the haemodynamic response of auditory cortex to time scanning of each trial to correspond to the peak BOLD response to a given tone. Long intertrial intervals ensure that any BOLD response related to loud scanner noises concomitant with scanning will have diminished before the next scan takes place.

Therefore, on each trial, a trigger from the scanner was sent to the stimulus presentation computer in order to synchronize timing between the two machines, so that the second test tone was presented 3.5 seconds before scanning onset (See Figure 1). Thus BOLD response to the critical test tone should be maximal during scanning (Belin et al., 1999). The 12.4 second inter-trial interval minimizes activation due to acoustic noise of the scanner.

In order to verify that the activation seen in our test scans was due to the test tones and not contextual tones, we performed a control condition using the random tone series. As in the test conditions, 6 seconds of contextual tones began 10.5 seconds before scanning, but no test tones were presented (See Figure 1). So instead of hearing two test tones at the end of the random tone series, there was silence. Subjects were instructed to listen to the tones, but that no pitch discrimination was necessary as no test tones would be presented. Since silence was heard in place of the test tones, we expected no difference between this condition and a silent comparison condition.

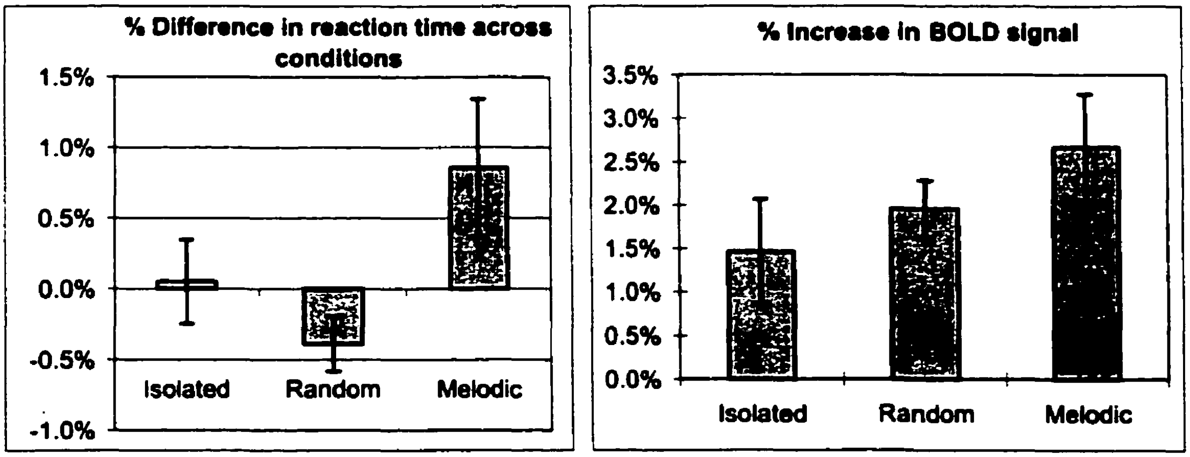

Figure 2: Chart on left shows average differences in reaction time across conditions, calculated as percent difference from individual medians (see text for details). Chart on right shows percent increases in BOLD signal in the right anterior STG compared to baseline (see text for details).

Results

Behavioral Results: Reaction time data were collected from 8 of the 9 subjects. Accuracy was generally very high, averaging 95%, 90%, and 90% correct in isolated, random, and melodic conditions respectively. Incorrect trials were not included in reaction time (RT) data. RTs were logged to fix rightward skewness of the results, and outliers were trimmed at >2.5 standard deviations, trimming 0 to 2 trials within each condition. RT results were standardized across subjects by calculating what percent faster or slower each condition was performed compared to all of that subject's trials using median values ((Overall RT-Condition RT)/Overall RT). Higher scores indicate faster performance. Although the effect of context was not significant as measured with an ANOVA ($F(2,14)=2.33$, $p>0.1$), the melodic context was performed fastest, as predicted, at 0.86% faster than overall. Isolated and random conditions did not differ from each other at 0.05% faster and -0.39% slower than overall, respectively ($p>0.05$) (See Figure 2). The melodic condition was performed significantly faster than the random condition ($t=2.01$, $p<0.05$, one-tailed), but not the isolated condition ($p>0.05$).

Control Results: The control condition confirmed that BOLD activity seen in the task conditions is due to the perception and processing of the test tones, not contextual tones. As predicted, when following the same time sequence as the task conditions, if no test tones are presented, resulting BOLD response differs only slightly from silence trials. One small area of activity was seen in the right lateral STG, with a negative peak in the left parietal region (See table 1).

Baseline Subtractions: Task conditions compared to baseline produced widespread activation, with many BOLD response peaks distributed throughout the cortex (See Table 2). However, these activations were confined to specific neural systems, and were consistent across tasks. For example, all temporal lobe activity occurred on the superior temporal gyri, areas known to be involved in auditory processing. Both STGs were nearly completely activated. All baseline subtractions also showed right frontal activity (See Figure 3). Motor and somatosensory regions were active, including pre- and post-central gyri, supplementary motor areas, and cerebellum, presumably due to pressing the mouse button. Other common activations included right insula and left inferior frontal regions. Negative peaks were not as consistent across baseline comparisons, but all included regions related to visual processing in both hemispheres (See Table 3).

Task / Task Subtractions: We performed three additional analyses comparing task conditions to each other: melodic minus isolated, random minus isolated, and melodic minus random. These comparisons confirmed the consistency of the baseline comparisons in that nearly all activity was subtracted out. In fact, the melodic minus random subtraction produced no positive or negative peaks (See Table 1).

The melodic minus isolated comparison yielded two positive peaks, meaning these areas were more active in the melodic than the isolated condition; the predicted right anterior STG (BA22), and right insula (See Figure 3). A weak co-activation of the left anterior STG was apparent, although well below significance (t=3.80). One negative peak was evident in medial precuneus. This area did not appear in either baseline subtraction, but other visual areas were seen as negative peaks in both conditions. Therefore, instead of interpreting this peak as a region of higher activity in the isolated condition, we suggest it reflects a larger decrease in activation in the melodic task. This negative precuneus peak was also observed in the random minus isolated comparison, and again we interpret it as a larger decrease in activation in the random condition. Two other negative peaks were found to be significant in this comparison, both in the anterior cingulate gyrus. No positive peaks reached significance, but a weak activation in the right anterior STG was apparent, suggesting slightly more activity in this region in the random than in the isolated condition.

Right Anterior STG: As the right anterior STG was the most critical region to emerge in the task/task subtractions, we investigated it further by performing analyses using individual subjects' scans. First, we calculated individual percent change in BOLD response of each test condition over baseline at the voxel showing peak activity in the right anterior STG in the group analysis: x=60, y=10, z=-6. These values increased with increasing context (See Figure 2), and a significant effect of context was found (F=4.985, df=2, p<0.05). We then performed paired sample one-tailed t-tests on these values, finding the percent change significantly higher in the melodic than in the isolated condition (t=4.74, p=0.0005), and the percent change in the random condition situated halfway between these conditions.

In order to assess the consistency of position of this anterior STG peak across individuals, we searched each subject's melodic vs. isolated comparison and located the STG peak closest to that seen in the group analysis. The region of interest was defined as the anterior aspect of the right STG, and if a discrete group of voxels was active in this region, the coordinates of the voxel displaying the highest t-value were recorded. A peak was discernable in 8 of the 9 subjects' scans, averaging 5.9 mm away from the group peak (range: 1.5-12.0 mm), indicating relatively high consistency across subjects.

Figure 3: (See previous page) Averaged fMRI subtraction images are shown superimposed upon an averaged anatomical MRI scan for 9 subjects. X and Z coordinates refer to standardized Talairach space. In all transverse sections, the right side of the brain is located on the right side of the figure. Color code for range of t-values is shown separately for each panel. **A:** Areas of increased BOLD activity in task conditions compared to baseline. The tail of the dotted yellow arrow in the top left image indicates the plane of section for the second row of images (z=-6). The tail of the solid red arrow indicates the plane of section for the third row of images (z=6). The arrow tails in the leftmost images of the second and third rows indicate the plane of section for the first row of images (x=50). All three task conditions produced extensive bilateral activity in the STG, as well as right dorsolateral frontal regions. Levels of BOLD activity were consistent across conditions in all areas except the right anterior STG which increased with increasing levels of context. The top three images (x=50) illustrate right STG and frontal activity. T-values in the anterior STG, indicated by the yellow arrow in the leftmost image, are seen to increase with increasing context. The second row of images (z=-6) illustrate the bilateral nature of anterior STG activity. Notice how the activity shown on the left stays constant across conditions whereas activity in right anterior STG, indicated by the arrow, increases with increasing context. The third row of images (z=6) illustrates the bilateral activity of the entire STG, as well as how this activity remains constant across conditions. **B:** Increased activity in melodic as compared to the isolated condition. The tail of the arrow in each image indicates plane of section of the other. Both arrows point to activity in the right anterior STG. Notice how nearly all of the BOLD activity shown in panel A has been cancelled out by this subtraction.

Comparisons Between Conditions			
Structure	BA	Coordinates	T
Control - Silence			
Positive Peaks			
R lat STG	22	(68,-30, 0)	5.15
Negative Peaks			
Med post par	7	(0,-60, 40)	-4.90
Melody - Isolated			
Positive Peaks			
R STG (Ant.)	22	(60, 10, -6)	5.64
L STG (Ant.)*	22	(-51, 3,-10)	3.80*
R Inferior Insula		(40, 2,-12)	4.96
Negative Peaks			
Med. Precuneus	7	(0,-64, 40)	-4.88
Random - Isolated			
Positive Peaks			
R STG (Ant.)*	22	(58, 14, -8)	3.56*
Negative Peaks			
Med. Ant.Cingulate	32	(0, 46, 4)	-5.62
L Ant. Cingulate	32/10	(-10, 46, 8)	-5.08
Med. Precuneus	7	(0,-56, 40)	-4.86
Melody - Random			
No Significant Positive or Negative Peaks			

Table 1: Activation foci for control condition and comparisons between conditions. X,Y,Z coordinates refer to Talairach stereotaxic space. T-values indicate statistical significance. Abbreviations: BA=Brodmann's area; L=left; R=right; Lat=lateral; Ant=anterior; Med=medial; STG=superior temporal gyrus; post par=posterior parietal. *Indicates t-values below significance levels.

Positive Peaks		**Isolated - Silence**		**Random - Silence**		**Melodic - Silence**	
Structure	BA	Coordinates	T	Coordinates	T	Coordinates	T
Temporal Lobe							
R Heschl's g. (A1)	41	(46,-18, 8)	5.15	(46,-16, 12)	5.45	(50,-14, 6)	5.56
R Heschl's g. (A1)	41					(42,-30, 18)	4.96
R Pl. Temp. (A2)	42	(62,-24, 14)	6.47	(62,-24, 14)	7.00	(62,-24, 14)	6.21
R Pl. Temp. (A2)	42	(62,-10, 6)	5.02	(40,-32, 16)	5.54	(62,-10, 6)	5.29
R Pl. Temp. (A2)	42	(40,-32, 16)	5.78	(62,-10, 10)	5.26		
R Ant. STG	22	(58, 6, -6)	5.07				
R STG	22	(46, -6, 6)	5.46			(46, -4, 4)	5.73
R STS	22					(66,-22, -2)	5.22
R STS	21/22	(60,-26, -2)	5.15	(60,-26, -2)	5.37		
R Temporal Pole	38			(50, 14, -6)	6.64	(52, 14, -6)	7.14
L Heschl's g. (A1)	41	(-44,-28, 10)	6.86				
L Pl. Temp. (A2)	42			(-64,-12, 8)	9.05		
L STG	22	(-62,-12, 8)	7.94			(-62,-12, 8)	7.87
L Ant. STG	22	(-56, 6, -4)	7.60	(-58, 8, -4)	6.78	(-56, 4, -4)	6.57
L Post. STG	22			(-44,-34,12)	5.94	(-44,-34, 12)	5.55
Frontal Lobe							
R Mid. Fr. g.	6			(44, 4, 52)	4.84		
R Mid. Fr. g.	9	(48, 16, 32)	6.82				
R Mid. Fr. g.	8/9					(48, 14, 36)	6.05
R Mid. Fr. g.	44/45			(50, 20, 24)	6.66		
R Inf. Fr. g.	47	(46, 14, 0)	6.02	(50, 16, 0)	6.78		
R Precent. g.	6					(44, 2, 52)	5.10
L Med. Fr. g.(SMA)	6	(-2, 14, 46)	8.57	(-4, 14, 46)	8.08	(-4, 14, 46)	8.80
L Med. Fr. g.(SMA)	6	(-4, -4, 54)	7.94	(-4, -2, 58)	7.62	(-4, -6, 60)	8.60
L Med. Fr. g.	32	(-2, 24, 34)	5.53	(-2, 24, 34)	4.81		
L Mid. Fr. g.	6	(-48, 4, 38)	6.43			(-48, 6, 38)	6.30
L Precent. g.	6			(-30,-10, 60)	4.82	(-38. -6, 38)	4.81
L Precent. g.	6					(-32,-12, 60)	5.35
L Precent. g.	4					(-32,-24, 58)	6.16
L Precent. s.	6			(-50, 6, 38)	6.59		
L Inf. Fr. g.	47	(-36, 20, 6)	5.50	(-40, 18, 4)	5.25	(-40, 18, 2)	6.02
Parietal and Occipital Lobes							
R Sup. Par. g.	7	(28,-46, 40)	5.53				
L Postcent. g.	1/ 2	(-56,-26, 40)	4.91	(-40,-34, 58)	6.16		
L Postcent. g.	4	(-38,-26, 56)	6.13				
L Sup. Par. g.	7	(-34,-46, 44)	5.43			(-34,-46, 44)	5.56
L Inf. Par. g.	40			(-44,-36, 42)	5.26	(-44,-34, 42)	5.12
L Lingual Gyrus	17/18	(-20,-74, 4)	4.81	(-10,-90, 4)	5.02		
Other							
R Insula				(42, -6, 6)	5.47	(40, -4,-10)	6.84
R Insula (Ant.)		(30, 20, 2)	6.77	(32, 20, 4)	6.84	(32, 20, 0)	6.78
R Putamen						(18, 4, 4)	4.82
R Cerebellum				(38,-54,-28)	4.81		
L Cingulate g.	23	(-4,-40, 20)	4.84			(-4, 24, 34)	5.63
L Cingulate g.	23	(-4,-26, 26)	5.11				
L Cerebellum		(-4,-48,-20)	5.24	(-40,-60,-28)	6.48	(-42,-60,-28)	5.75
L Thalamus						(-14,-18, 8)	5.01

Table 2: See following page for table caption.

Table 2: (See previous page) Positive activation foci for baseline comparisons. Coordinates refer to Talairach stereotaxic space (X,Y,Z). T-values indicate statistical significance. Abbreviations: BA=Brodmann's area; L=left; R=right; g=gyrus; s=sulcus; Pl Temp=planum temporale; Fr=frontal; Precent=precentral; Postcent=postcentral; Par=parietal; Occ=occipital; Ant=anterior; Post=posterior; Sup=superior; Inf=inferior; Mid=middle; Med=medial; Lat=lateral; STG=superior temporal gyrus; SMA=supplementary motor area.

Negative Peaks		**Isolated - Silence**		**Random - Silence**		**Melody - Silence**	
Structure	BA	Coordinates	T	Coordinates	T	Coordinates	T
Temporal Lobe							
R Mid. Temp. g. (post.)	39			(42,-68, 26)	6.19		
R Mid. Temp. s.	20/21					(54,-12,-18)	5.17
R Inf. Temp. g.	20			(54,-12,-18)	5.05		
L Inf. Temp. g. (post.)	19					(-50,-80, 12)	4.81
Frontal Lobe							
R Sup. Fr. g.	8	(22, 24, 44)	4.90	(22, 24, 44)	6.14		
R Mid. Fr. g.	8					(26, 30, 46)	5.31
L Mid. Fr. g.	8			(-26, 22, 42)	5.45	(-26, 22, 42)	5.97
Parietal Lobe							
L Precuneus	7			(-6,-54, 40)	5.74		
Occipital Lobe							
R Sup. Occ. g.	19			(34,-78, 26)	5.10	(38,-80, 24)	5.55
R Mid/Sup. Occ. g.	19	(42,-74, 26)	5.95				
L Mid/Sup. Occ. g.	19	(-44,-74, 24)	6.07	(-44,-72, 24)	5.71		
L Mid. Occ. g.	39/19					(-46,-74, 26)	5.56
L Fusiform g.	37	(-30,-42,-12)	5.07				
Other							
R Post. Cingulate g.	31			(6,-56, 24)	5.20		
R Ant. Cingulate g.	32			(2, 44, -2)	5.53	(2, 42, -2)	4.87
L Ant. Cingulate g.	32			(-14, 42, 10)	5.21		

Table 3: Negative activation foci for baseline comparisons. Coordinates refer to Talairach stereotaxic space (X,Y,Z). T-values indicate statistical significance. Abbreviations: see caption for Table 2.

Discussion

In general, the predictions outlined above were fulfilled. Both the right anterior STG and right frontal regions were involved in the pitch discrimination tasks performed under different contextual conditions. Most interestingly, right anterior STG showed increasing activity with increasing levels of context. This was seen in the task / task comparisons in the group data, as well as individual subjects' BOLD response levels in this region from task scans compared to baseline. The increased activation of this region with increasing levels of context appears to be relatively consistent across subjects, with all but one subject's melodic minus isolated comparison showing a discernable activation within millimeters of the peak seen in the group data.

Although right dorsolateral frontal activity did not increase with increasing context, this region was active in each condition when compared to baseline activity, presumably due to the working memory aspect of the pitch discrimination task common to all conditions. The consistent activation level of this region across conditions argues that when comparing the pitch of two consecutive tones, the amount of processing done in right frontal areas is similar across differing levels of context.

The near complete activation of bilateral STG in the baseline comparisons is consistent with other studies using tonal stimuli (Zatorre et al., 1996; Warrier et al., 1999b). This type of auditory stimulus requires much more processing than simple frequency discrimination of pure tones, thus recruiting many different auditory cortical areas, including primary, secondary, and association regions. Another commonality with previous auditory imaging studies is that decreased activity is often reported in regions associated with visual processing when listening to tones is compared to a resting state (Holcomb et al., 1998; Zatorre et al., 1999). Although the exact nature of these decreases is unknown, perhaps there is more visual processing during baseline than in task conditions. During silent trials, subjects may be generally less focused, distributing sensory processing more evenly across sensory systems. However, when focusing on a particular modality, audition for example, relative increases are seen in auditory areas, while lack of attention to other modalities causes relatively less processing in regions associated with those systems.

Melodic context produced the most facilitation behaviorally, seen as faster reaction times, and also produced the most activity in right anterior STG. This pattern of behavioral and functional results fits well with our idea that auditory processing in the right anterior STG contributes to the effect of contextual tones. While individuals' BOLD signal in the functional scans increased with increasing tonal context, not all subjects showed the same linearly increasing behavioral scores. However, most showed the predicted facilitation with melodic as compared to isolated context. Information held in the random series of tones did not facilitate tone discrimination in this study, and results from this and the isolated condition did not differ. A previous study including a tonal context other than isolated and melodic had used a non-random predictable tone series, producing behavioral results intermediate to that of the other two (Warrier & Zatorre, Submitted). Pilot testing with the random condition had indicated similar results, although this did not occur reliably in the final testing. However, whatever the pattern of results, most subjects showed marked behavioral differences between conditions, and as melodic context produced the quickest behavioral results as well as the highest activation in right anterior STG, we suggest this region plays a facilitating role in incorporating contextual information into pitch discrimination. This auditory association area has cortico-cortical connections to right frontal regions implicated in tonal working memory, giving the region a possible mechanism for processing tones relative to those previously heard.

We cannot say from these results that changes in anterior STG activity cause the behavioral differences seen between conditions. We have merely observed a relationship between the behavioral and functional results. Other methods of investigation are needed to address this question. That said, the result fits well with a lesion study in which neurosurgical patients with right anterior temporal lobe removals profited much less from the presence of melodic context on a pitch discrimination task than patients with similar removals on the left or a control group(Warrier & Zatorre, 1999). The patients' lesions included more cortex than just the anterior aspect of the STG. However, all excisions included the area seen in this study to relate to processing tones relative to contextual tones. In addition, posterior extent of the lesions (encroachment onto Heschl's gyrus) did not affect outcome on the task, implicating anterior regions in the critical processing.

The present results combined with those of the lesion study thus make a stronger case for involvement of the right anterior STG in tonal contextual effects.

The control condition showed that listening to the contextual tones produced minimal BOLD response, contrasting greatly with the task conditions. The one region showing this response was located on the right posterior STG, an area previously shown to be involved in processing tonal stimuli (Zatorre et al., 1994). This area is distinct from the right anterior STG region seen in the task conditions, located over 4cm away. Therefore, the control condition verified that the activity of regions implicated by the task conditions in producing contextual effects is due to processing of the test tones, and not the contextual tones.

Conclusion

The main finding of this study is that right anterior STG appears to have a role in producing tonal context effects. This fits in well with data from a previous lesion study finding patients with resections of this area to have impaired use of contextual cues in a similar tone discrimination task. This region's connections to right frontal cortex, involved in tonal working memory, appear to allow processing of tones relative to those heard previously. Activation of right anterior STG increased significantly with each level of increased tonal context. Melodic context generated the most activity in this region, and produced the quickest pitch discrimination behaviorally, suggesting that activity in this region facilitates pitch discrimination.

Chapter 5

Investigating the Separation of Lyrics and Melody in Song through PET

Listening to someone singing can be a beautiful auditory experience. But what comes to our ears is merely a complex pattern of fluctuations in air pressure. How does our brain take these fluctuations and decode them into the combined percept of lyrics and melody? Are they processed and encoded as a unified sound object, or are the lyrics and melody processed separately? When examined in isolation, i.e. spoken text or a melodic line, each appears to be processed by different neural systems (e.g. Zatorre et al., 1994; Demb et al., 1995). But what happens when the text and melody are fused in time as they are in song? Does the brain pull apart the text and the melody and process them separately, each within its respective neural system, or is song a "special case" that gets processed by a system of its own? This study will examine the issue from a functional neuroimaging perspective. Using sung opera excerpts, we will determine whether or not processing of the melodies and lyrics of these excerpts occurs in different neural systems. We will begin by reviewing some of the research arguing for and against the separate processing of lyrics and melody in song, then move on to describe the present study and discuss how it contributes to this debate.

Serafine, Crowder and Repp investigated the topic by performing a series of recognition experiments with relatively unknown American folk songs. Their results support the theory that song is a special case in auditory perception in that the lyrics and melody of a song are integrated in memory (Serafine et al., 1984; Serafine et al., 1986; Crowder et al., 1990). In their first experiment, the lyrics and melodies of the songs initially presented for encoding (old songs) were more easily recognized than those of "mismatch" songs, (lyrics sung to melodies previously paired with different lyrics). This effect held when subjects were instructed to attend to the melodies only (Exp. 2, 1984), and when the melodies were sung with nonsense text (1986, 1990).

Previous cognitive psychology studies had found evidence of a separation of musical and verbal dimensions in memory. A useful method for this sort of research is the interpolated sequence paradigm, in which a listener determines if the first and last events of a sequence are the same or not. Qualities of the intervening events are varied systematically. If they cause decreased recognition of the test events, the dimensions of the test and intervening events are inferred to be integrated in memory in some way. If, however, no interference in recognition is observed by the intervening events, the dimensions are inferred to be separate in memory. Using this paradigm, memory for pitch and spoken digits was found to be separable (Deutsch, 1970). The pitch of the interpolated events interfered with pitch but not digit memory, and the spoken digits interfered with digit but not pitch memory. However, in contrast to the sung lyrics used in Serafine et al's studies, the verbal and musical dimensions were contained in separate events, not fused, as in song. But dichotic listening studies that get closer to this song format through sung letter or number strings also find the musical and verbal dimensions to be separable (Bartholomeus, 1974; Goodglass & Calderon, 1977). In these studies, two letter or number strings sung to different melodies were presented simultaneously, one to each ear. A left ear superiority was found when subjects attended to the melody, while a right ear superiority was found when subjects attended to the letters and numbers. The opposing laterality effects seen in these studies argue for the separate processing of musical and verbal material.

The integration theory was again challenged by Samson and Zatorre (1991) when, using Serafine et al's folksong stimuli (1984), they found more evidence for separation than integration of lyrics and melodies when testing three groups of listeners: patients with right temporal lobe lesions (RT), patients with left temporal lobe lesions (LT), and a neurologically normal control group. In support of the integration theory, melodies were better remembered with old lyrics than with new ones in the control group. However, they did not remember the lyrics better with old melodies, suggesting separation of these dimensions. This result could not be attributed to a ceiling effect as it had been in Serafine et al's study. Differential performance by the two patient groups also suggested separation of lyrics and melody in memory. For example, the LT group was impaired overall in both lyric and melody recognition compared to the control and RT groups. In addition, while only the LT group showed impairments on the old lyric/new melody condition, both LT and RT groups performed near chance in the old melody/new lyric condition. When the stimuli consisted of melodies without lyrics (sung with "la" in place of text), or lyrics without melodies (spoken words), a double dissociation was found such that the LT group was impaired on word but not melody recognition, while the RT group was impaired on melody but not word recognition. These differences in performance between the two patient groups suggest that the different dimensions of song stimuli are processed in different cortical areas, thus supporting the separation theory.

Another line of evidence pointing toward the separate encoding of lyric and melody in song comes from amusic patients, people with neurological damage resulting in a disruption of musical abilities while sparing the ability to perform language abilities. Patients CN and GL, who have bilateral superior temporal cortex damage, show functional dissociations when perceiving and recognizing sounds, such that linguistic and environmental sound perception is spared while perception of tunes, prosody, and voice is impaired (Peretz et al., 1994). For example, CN easily recognizes lyrics from familiar songs, but cannot recognize the melodies of these same songs. Similarly, GL's recognition of familiar lyrics is well above chance, while his melody recognition of the same songs falls far below that of matched control subjects. Patients with the reverse symptomology have also been documented (e.g. Luria et al., 1965) showing a double dissociation between memory for song lyrics and memory for the melodies of these same songs. This dissociation in the patient literature argues for separate encoding of the lyrics and melody of songs.

An issue intimately related to whether or not lyrics and melody are stored separately in the brain, is whether or not they are processed separately during their initial neural encoding. A recent electrophysiological investigation into this matter found evidence that song lyrics and tunes are processed separately (Besson et al., 1998). In this study, subjects listened to *a capella* opera excerpts, i.e. excerpts sung without instruments, in which the final word could be semantically correct or incongruous, and in which the final note could be sung in or out of tune. The task was to determine if the end event of each excerpt was congruous or incongruous with respect to the words and/or the melody, while scalp event-related potentials were recorded. In accordance with previous verbal studies, Besson and colleagues saw N400 responses to semantic incongruities (e.g. (Kutas & Hillyard, 1980). A late positive component (LPC) was seen with melodic incongruities, also in accord with prior studies of musical processing (Besson & Faita, 1995; Janata, 1995; Patel et al., 1998). When they presented both semantic and melodic incongruities simultaneously, this elicited both the N400 and LPC components. The waveform to this double-incongruous condition was well-predicted by adding the waveforms from semantic and melodic incongruities separately, suggesting separate neural generators for each waveform.

So is song a "special case"? There appears to be more evidence pointing to the separation of song lyrics and music in memory than to their integration. However, Serafine and colleagues have replicated their results, finding them robust under a variety of manipulations, and therefore should not be ignored. The present study aims to use PET methodology to help determine whether or not lyrics are processed independently from the melody in song using a paradigm based on that of Besson et al (1998). Using their stimuli consisting of excerpts of opera sung *a capella*, our subjects determined if the last event of each excerpt was semantically or melodically congruous or incongruous with the rest of the excerpt. During each scan, they either attended to the lyrics or the melodies. This manipulation allowed us to examine possible differences in brain activity due to directing attention to different aspects of the stimuli. Table 1 outlines the six conditions. Figure 1 shows a sample of one opera excerpt with the four possible endings indicated.

Test Conditions		
	MONITOR SEMANTICS	MONITOR MELODIES
CC	*Condition Sem_CC:* Semantically Congruous/ Melodically Congruous	*Condition Mel_CC:* Semantically Congruous/ Melodically Congruous
IC	*Condition Sem_IC:* Semantically Incongruous Melodically Congruous	*Condition Mel_IC:* Semantically Congruous/ Melodically Incongruous
II	*Condition Sem_II:* Semantically Incongruous/ Melodically Incongruous	*Condition Mel_II:* Semantically Incongruous/ Melodically Incongruous

Table 1: Descriptions of test conditions

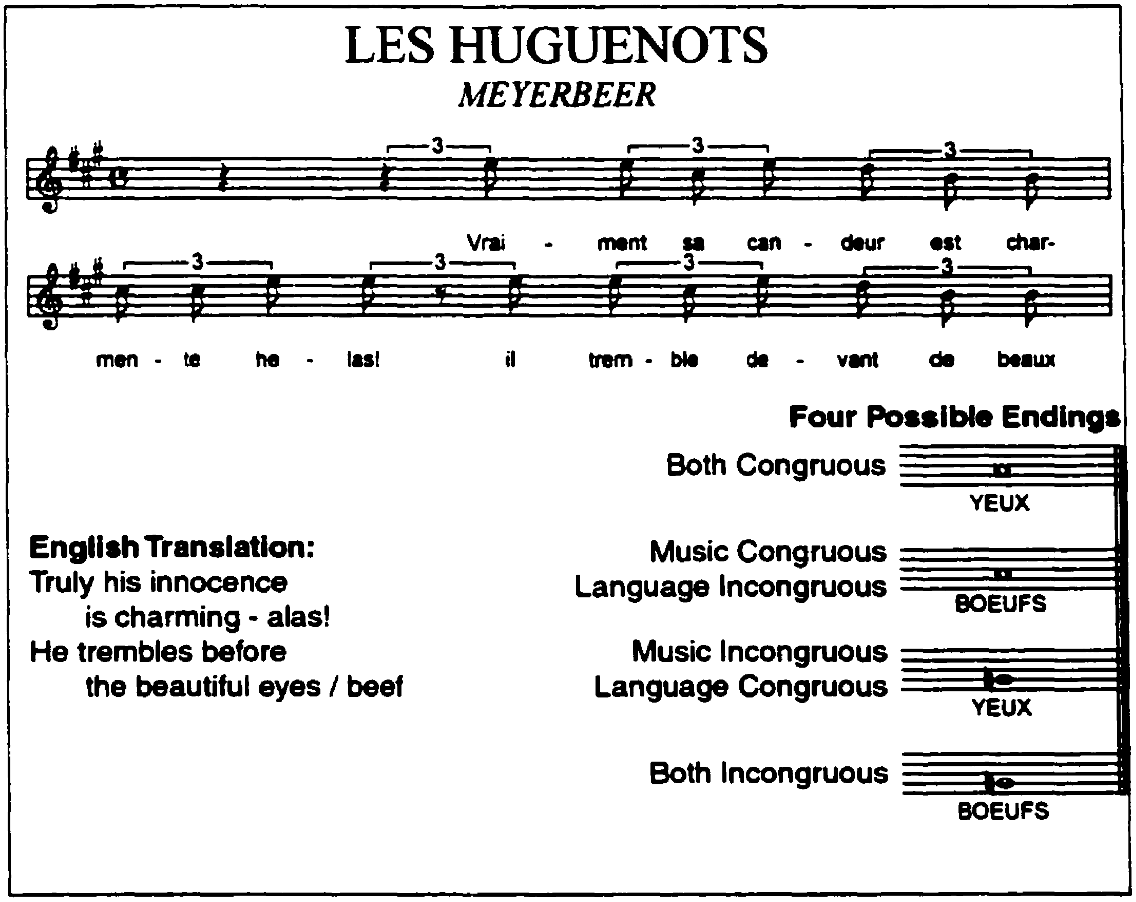

Figure 1: Sample stimulus shown in musical notation. Condition abbreviations follow Table 1.

If song is a special case, and the brain processes lyrics and melody as a unified percept, the same pattern of activation should be apparent when subjects monitor for semantic or melodic incongruities. However, if they are processed separately, different patterns of activation should be evident during semantic and melodic monitoring. Basing our hypothesis on previous lesion and electrophysiological research, we predicted different patterns of activation when subjects monitored lyrics versus melody. In particular, when monitoring for semantic incongruities we predicted the pattern of activation to include left inferior frontal cortex, an area consistently found in imaging studies in which subjects perform various semantic tasks (Petersen et al., 1988; Binder et al., 1995; Demb et al., 1995; Stromswold et al., 1996; Vandenberghe et al., 1996; Binder et al., 1997; Klein et al., 1997; Bookheimer et al., 1998; Klein et al., 1999). When monitoring for melodic incongruities, we predicted right-asymmetric frontal and temporal cortical activations (Zatorre et al., 1992; Zatorre et al., 1994; Binder et al., 1997; Griffiths et al., 1999; Perry et al., 1999).

We were also interested in investigating which neural structures were responsible for the N400 and LPC evoked potentials seen in Besson et al's study with incongruous trials (1998). In that study it was evident that these two potentials reflected separate, parallel processes. This suggests that they are generated by different brain structures, but the EEG methodology used in that study is unable to localize exactly which structures are contributing to the potentials. Since the PET methodology is better suited to this sort of localization, we constructed the conditions in such a way as to be able to look directly at this problem. This was achieved by testing congruous and incongruous stimuli in separate conditions in order to be able to compare them. Although this was an exploratory comparison, we broadly predicted that the detection of incongruous stimuli, when compared to the processing of congruous stimuli, would result in increased left-sided activation within language areas, such as the LIF or posterior temporal regions when the semantics were incongruous, and that areas related to music processing in the right temporal and right frontal regions would be recruited when melodic incongruities were detected.

Results

Behavioral Data: Behavioral responses indicating whether the relevant dimension was congruous or not were recorded. All subjects performed well on this measure, with an average of 93.8% correct overall (range: 90.7% - 98.1%). Using a paired samples t-test, we determined that more errors occurred when attending to the semantics (average = 2.09) than when attending to the melodies (average = 1.27) t (one-tailed) = 1.845, $P<0.05$. This resulted in 92.3 and 95.3% correct for semantics and melodies, respectively.

Functional Data: Three sets of analyses were performed on these data. In the first analysis, we simply compared each test condition to a noise baseline condition. This was intended to isolate structures whose activation reflects processing over and above simple auditory processing. In the second analysis we examined the effects of attention by comparing conditions in which listeners attended to the semantics to those in which they attended to the melodies. Two subtractions were performed in this analysis. The first compared the two conditions in which both semantics and melodies were congruous, and the second pooled all semantic monitoring conditions, and all melodic monitoring conditions, and compared these results to one another. The third analysis examined congruity effects both when monitoring the melodies and monitoring the semantics. Each of these effects was assessed by subtracting the double congruous condition from the condition in which the relevant dimension only was incongruous. A list of all comparisons is contained in Table 2.

STATISTICAL COMPARISONS

Baseline Comparisons:	Congruity Effects:
Sem_CC - Baseline	Sem_IC - Sem_CC
Sem_IC - Baseline	Sem_CC - Sem_IC
Sem_II - Baseline	Mel_IC - Sem_CC
Mel_CC - Baseline	Mel_CC - Mel_IC
Mel_IC - Baseline	
Mel_II - Baseline	
Attention Effects:	
Sem_CC - Mel_CC	
Mel_CC - Sem_CC	
(Sem_CC + Sem_IC + Sem_II) - (Mel_CC + Mel_IC + Mel_II)	
(Mel_CC + Mel_IC + Mel_II) - (Sem_CC + Sem_IC + Sem_II)	

Table 2: Statistical comparisons performed. All comparisons are shown as subtractions. Condition abbreviations follow Table 1.

Baseline subtractions: Every test condition minus the noise baseline showed pervasive bilateral activity of the auditory cortices throughout the superior temporal gyrus (STG) as expected (See Figure 2-A). The peak of this activity in the left STG in each scan was located in the middle of the gyrus, BA22. Each test scan also produced two peaks on the right STG, one near the middle of the gyrus, BA22, and one more anterior, bordering BA's 22 and 38 (See Table 3). Motor areas such as the precentral gyrus and cerebellum were also activated in these subtractions. However, as they most likely pertain to the motor response, not the cognitive functions involved in performing the task, and are common to all test conditions, these regions will not be discussed. Interested readers will find this information reported in Table 3.

As predicted, monitoring for semantic incongruities produced activation localized to the left inferior frontal cortex (LIF; areas 44 and 45) (See Figure 2-A). This result was robust, occurring in all baseline subtractions from scans monitoring semantics ($t > 4.0$ for all). Melodic monitoring did not produce activation in LIF. No cortical areas showed selective activation during melodic monitoring in these subtractions.

Five out of the six test conditions, when compared to the noise baseline, showed insular activation in either the right or the left hemisphere. No pattern of laterality could be discerned. However, orbital frontal activation occurred in the same five test conditions, and was always seen on the same side as the insular activation. Two conditions, semantic monitoring with double congruity (Sem_II), and melodic monitoring with double incongruity (Mel_II), showed bilateral orbital frontal activation. Right anterior cingulate activation (BA32) was seen in the double incongruous monitoring for semantics condition.

Attentional effects: Different sets of structures were activated when subjects monitored for semantic incongruities than when they monitored for melodic incongruities. Different activation patterns emerged even when the stimuli were exactly the same, as when comparing semantic vs. melodic monitoring in the two conditions in which no incongruities were present, our first attentional comparison (Sem_CC vs. Mel_CC) (See Table 4). In this comparison, monitoring for semantic incongruities produced relatively more activation in the LIF, in addition to a right orbital frontal area. Monitoring for melodic incongruities, on the other hand, produced relatively more activity in the left superior parietal lobe (BA7), and the right posterior cingulate (BA23).

In order to increase our statistical power, we averaged all three scans monitoring semantics and compared them to an average of the three scans monitoring melodies (Sem_CC + Sem_IC + Sem_II vs. Mel_CC + Mel_IC + Mel_II). This increased the number of scans involved in the subtraction from 22 to 66. We again found different patterns of activation when monitoring different aspects of the stimuli (See Table 4 and Figure 2-B). Two LIF areas were more active when monitoring for semantic incongruities (BA44 and 45), one of which peaked within millimeters of the LIF peak seen in the Sem_CC – Mel_CC comparison (See Table 4 for coordinates). The left posterior STG, was also more active in the semantic conditions. Like the LIF, this area is known to be involved in language processing. Although it was not seen in comparisons involving pairs of scans, this area reaches significance (t = 4.2) when all scans are pooled to increase the signal to noise ratio. A right orbital frontal area was also significant. An area in the right middle temporal gyrus, similar to the posterior temporal activation seen on the left when attending semantics, was more active when monitoring melodies. As can be seen in Figure 2-B, this area of activation is at the most posterior end of the middle temporal gyrus, near the temporo-parieto-occipital junction.

Congruity effects: We performed two subtractions examining the effects of congruity: 1) Monitoring semantics with a double congruity vs. semantics only incongruous (Sem_CC vs. Sem_IC), and 2) Monitoring melodies with a double congruity vs. melodies only incongruous (Mel_CC vs. Mel_IC). The effects of congruity were not easy to interpret. When monitoring for semantics, four areas were more active in the double congruous condition; right and left orbital frontal areas, and right and left cerebellar areas. Two areas were more active in the incongruous scan; the left superior parietal lobe (BA7), and a right frontal premotor area (BA6). No structures were more active in the congruous scan when monitoring the melodies, and only the right orbital frontal area was more active in the incongruous scan (See Table 5).

Figure 2: (On previous page.) Averaged PET subtraction images are shown superimposed upon the averaged MRI scan for the 11 subjects tested. X and Z coordinates refer to standardized Talairach space. In all horizontal figures, right side of the brain is located on the right side of the figure. Color code for range of t-values is shown in each panel. A: Areas of increased CBF activity in test conditions compared to baseline. Semantic monitoring conditions located on left half of panel, melodic monitoring conditions located on the right. Saggital and horizontal slices are shown for each condition, centering on peak of LIF activation in semantic monitoring conditions, with corresponding slices shown in melodic monitoring conditions. Condition abbreviations follow Table 1. Arrows indicate LIF foci. B: Areas of increased CBF activity related to attention in combined comparison (all three semantic monitoring conditions averaged together and compared to average of all three melodic monitoring conditions). Saggital and horizontal slices are shown centered on regions of interest. The upper three figures correspond to increased activity in semantic monitoring, while the bottom pair corresponds to increased activity in melodic monitoring. In the upper saggital slice (x = -50), activation foci are color-coded such that green and red arrows point to separate LIF peaks, while the yellow arrow indicates posterior STG activity. Corresponding horizontal slices (z = 24 and z = 6) maintain this color code.

Table 3: (On following page.) Activation foci for all baseline comparisons. X,Y,Z coordinates refer to Talairach stereotaxic space. T-values indicate statistical significance. Abbreviations: L=Left; R=Right; SMA=Supplementary Motor Area; BA=Brodmann's Area; CC=Both dimensions congruous; IC=Relevant dimension incongruous; II=Both dimensions incongruous. *Nomenclature following Petrides and Pandya, 1994.

Test Conditions compared to Baseline

		Monitor Semantics		Monitor Melodies	
Substrate	Cond	(x, y, z)	t	(x, y, z)	t
L. Superior Temporal Gyrus (BA22)	CC	(-60,-21, 6)	12.3	(-60,-19, 6)	11.8
	CC			(-52, 5, -6)	8.7
	IC	(-59,-18, 5)	14.2	(-59,-16, 5)	12.5
	II	(-59,-18, 6)	15.0	(-59,-19, 6)	13.9
R. Superior Temporal Gyrus (BA22)	CC	(59,-11, 2)	10.0	(59,-13, 2)	10.4
	IC	(59, -9, 2)	11.5	(59, -9, 11)	10.7
	II	(59, -9, 2)	13.1	(59, -9, 2)	11.7
R. Superior Temporal Gyrus (BA 22/38)	CC	(8,-11, 8)	8.1	(54, 6, -9)	8.0
	IC	(52, 10,-11)	9.1	(54, 12,-14)	7.7
	II	(52, 10,-11)	8.8	(55, 10,-11)	7.5
L. Inferior Frontal (BA45, BA47)	CC	(-44, 20, 9)	4.1	-	-
	IC	(-48, 24, 14)	4.7	-	-
	II	(-50, 22, 14)	5.7	-	-
R. Anterior Cingulate (BA32)	II	(3, 32, 27)	4.3	-	-
L. Anterior Insula (borders on BA47)	CC	(-32, 22, 5)	3.6	-	-
	II	-	-	(-28, 22, 9)	4.1
R. Anterior Insula (borders on BA47)	IC	(29, 20, 9)	5.0	(34, 20,12)	4.2
	II	(34, 22, 6)	4.0	-	-
L. Orbital Frontal (Area 14*)	CC	(-1, 17,-18)	6.0	-	-
	II	-	-	(-5, 20,-18)	4.2
	II			(-3, 18,-17)	4.2
R. Orbital Frontal (Area 14*)	CC	(17, 20,-18)	5.6	-	-
	IC	(4, 13,-17)	3.9	(1, 15,-17)	5.0
	II	(15, 20,-18)	3.9	(15, 30,-14)	3.6
L. Precentral Gyrus (BA4)	CC	(-51, -7, 53)	4.0	(-48,-11, 54)	3.8
	IC	(-47, -9, 50)	4.5	(-50,-11, 50)	4.2
	II	(-47, -9, 53)	4.7	(-48, -9, 53)	4.2
L. Medial Frontal – SMA (BA6)	II	-		(-7, 10, 60)	3.7
L. Cerebellum	CC	(-29,-62,-21)	5.5	(-28,-52,-23)	3.3
	IC	(-24,-62,-15)	4.3	(-23.-59,-21)	4.6
	II	(-17,-59,-21)	4.0	-	
R. Cerebellum	CC	(34,-61,-21)	6.5	(4,-69,-20)	4.0
	CC	(9,-64,-18)	5.0	(30,-61,-23)	3.8
	IC	(15,-66,-23)	5.5	(32,-59,-23)	5.1
	IC	(29,-59,-20)	4.1	(9,-64,-21)	5.1
	II	(34,-61,-23)	5.6	(35,-61,-21)	4.2
	II	(11,-64,-23)	3.5		
R. Caudate Nucleus	CC	-		(17, 13, 8)	4.3
Brainstem	CC	-		(3,-23, -9)	3.7
	IC	(3,-26,-14)	4.4	-	
	II	(8,-26,-14)	4.6	-	

Table 3: (See previous page for caption.)

Attentional Effects

Substrate	Comparison	Monitor Semantics (x, y, z)	Monitor Semantics t	Monitor Melodies (x, y, z)	Monitor Melodies t
L. Inferior Frontal (BA44,45,47)	Sem_CC vs Mel_CC	(-46, 29, 0)	3.3	-	-
	Combined cond's	(-47, 30, 3)	4.2	-	-
	Combined cond's	(-50, 17, 24)	4.4		
L. Post. Sup. Temp. G. (BA22)	Sem_CC vs Mel_CC	-	-	-	-
	Combined cond's	(-55,-33, 6)	4.2	-	-
R. Post. Mid. Temp. G. (BA21)	Sem_CC vs Mel_CC	-	-	-	-
	Combined cond's	-	-	(50,-54, 15)	3.3
L. Superior Parietal (BA7)	Sem_CC vs Mel_CC	-	-	(-15,-62, 62)	4.4
	Sem_CC vs Mel_CC			(-27,-44, 45)	3.6
	Combined cond's	-	-	-	-
L. Anterior Cingulate (BA32)	Sem_CC vs Mel_CC	-	-	-	-
	Combined cond's	-	-	(-7, 37,-12)	3.6
R. Posterior Cingulate (BA23)	Sem_CC vs Mel_CC	-	-	(3,-35, 27)	3.5
	Combined cond's	-	-	(4,-33, 33)	3.6
R. Orbital Frontal (Area 14*)	Sem_CC vs Mel_CC	(15, 17,-17)	5.1	-	-
	Sem_CC vs Mel_CC	(4, 18,-20)	4.3		
	Combined cond's	(15, 15,-15)	3.5	-	-

Table 4: Activation foci of attentional comparisons. X,Y,Z coordinates refer to Talairach stereotaxic space. T-values indicate statistical significance. Abbreviations: L=Left; R=Right; Post.=Posterior; Sup.=Superior; Temp.=Temporal; Mid.=Middle; G.=Gyrus; BA=Brodmann's Area; Sem_CC = semantic monitoring when both dimensions are congruous. Mel_CC = melodic monitoring when both dimensions are congruous; Combined cond's = Average of all semantic monitoring conditions compared to average of all melodic monitoring conditions. *Nomenclature following Petrides and Pandya, 1994.

Congruity Effects

Substrate	Monitor	Congruous (x, y, z)	Congruous t	Incongruous (x, y, z)	Incongruous t
L. Superior Parietal (BA7)	Semantics	-	-	(-16,-61, 63)	3.8
R. Orbital Frontal (Area 14*)	Semantics	(19, 22,-20)	5.2	-	-
	Semantics	(1, 46,-15)	3.8		
	Melodies	-	-	(12, 12,-18)	4.3
L. Orbital Frontal (Area 14*)	Semantics	(-8, 22,-18)	6.0	-	-
R. Premotor (BA6)	Semantics	-	-	(13, -4, 59)	5.1
L. Cerebellum	Semantics	(-9,-69,-15)	4.1	-	-
R. Cerebellum	Semantics	(36,-62,-21)	4.0	-	-

Table 5. Activation foci of congruity comparisons. X,Y,Z coordinates refer to Talairach stereotaxic space. T-values indicate statistical significance. Abbreviations: L=Left; R=Right; BA=Brodmann's Area. *Nomenclature following Petrides and Pandya, 1994.

Discussion

Similarities across conditions:

Similarities were observed across all six test conditions compared to baseline, reflecting the common stimulus form used. All test conditions showed extensive bilateral STG activation (See Figure 2-A). This was not unexpected, because even though the baseline scan consisted of white noise matched in amplitude envelope to the stimuli, the opera excerpts are very rich aurally. They are sung with a human voice, follow complex melodic structures, and contain phonological information. These attributes of the stimuli require much more auditory processing than noise, as evidenced by the extensive activation of auditory cortex when activation to noise is subtracted out. When examining peaks within this extensive activity, we see two peaks on the right, and one on the left for most conditions. The differences may correspond to differential activation of auditory cortical areas on the right and the left, however, the same area on the left showed increased activation, although no particular peak could be discerned. We therefore suspect that this slight difference does not reflect functional differences between the two auditory cortices. Such broad bilateral activation of the STG has been seen in other studies using musical stimuli (e.g. Zatorre et al., 1996). This is in accord with the idea that in addition to the primary auditory cortices, other areas on the STG are involved in decoding / encoding various aspects of the acoustic signal (Brugge & Reale, 1985).

Orbital frontal and insular activation were also seen in all but one of the test condition minus baseline comparisons. Interestingly, the lateralization of insular activation seemed linked to orbital frontal cortex in that when insular activity was seen on the left, the left orbital frontal region was also seen, and when insular activity was seen on the right, right orbital frontal activity was also seen. Due to the caudal nature of the orbital frontal activations and the anterior position of the insular activations, this co-activity fits well with the anatomical connection literature. These two areas, both agranular regions, are reciprocally connected ipsilaterally (Mesulam & Mufson, 1982a,b), and are part of an orbital-insular-temporal circuit that also includes the agranular area of temporopolar cortex. We are unable to speculate at this time as to the exact function of these regions in our task.

Attentional Effects:

In contrast to these similarities across all test conditions, the predicted LIF area appeared in all conditions in which listeners attended to semantics, but none of the conditions in which they attended to the melodies, even though the stimuli in all conditions were the same (See Figure 2-A). Considered a language area, this region has consistently been shown to be involved in processing semantic aspects of language under a variety of tasks, such as noun classification (Binder et al., 1995; Binder et al., 1997), responsive naming (Bookheimer et al., 1998) synonym generation (Klein et al., 1997), verb generation (Petersen et al., 1988; Klein et al., 1999), semantic encoding (Demb et al., 1995), semantic plausibility judgments, (Stromswold et al., 1996), and associative semantic matching (Vandenberghe et al., 1996). It is interesting to note that in our study, this area was active only when listeners directed their attention toward the song lyrics, even though the lyrics were also heard when monitoring the melodies. This implies that simply hearing words is not sufficient for activating the LIF: a deeper processing of those words is necessary, requiring attention to be directed toward the words.

When looking at the same set of baseline comparisons, no structures were found to be active during melodic attention conditions only. The lack of right frontal activity in particular caught our attention. After seeing such strong and consistent left frontal activity when monitoring semantics, we might have expected at least weak activity in the contralateral frontal lobe when monitoring melodies. However, this non-result could be due to several different reasons. First and most obviously, there may be nothing to detect; perhaps all the necessary processing for this task is done in the auditory cortex. Alternately, we may not have enough power to be able to see an effect when inspecting single conditions, a problem we addressed in a separate analysis (see below). However, the single comparisons to baseline were sufficient to show activity in the conditions in which semantics were monitored. This suggests that the effect seen in the LIF is more robust, while an effect of melodic attention is either more variable, not as strong, or absent. Another factor that may be affecting effect strengths is difficulty. We attempted to avoid differences in difficulty between conditions by choosing only the easiest excerpts, hoping that attending to the lyrics and the melodies would be equally obvious.

But while out of key notes in the melodies are easily detected, attending to the semantics in these stimuli may be inherently more difficult. For instance, as in most opera texts, the language used in many of the excerpts is archaic, using words and phrases not typical of modern French. In addition, many times a sung phrase follows a different rhythm and emphasizes different syllables than if that phrase were spoken, requiring the listener to devote more attention to parsing the words. This difference in difficulty level is reflected in the behavioral data. Although overall performance was high, more errors were made when attending to the lyrics as compared to the melodies. Other studies have shown that as task difficulty increases, so does the activation related to that task (Baker et al., 1996; Winstein et al., 1997; Carpenter et al., 1999; Catalan et al., 1999). Therefore, slight differences in task difficulty may be contributing to the reason we see activation preferential to the semantic attention conditions, but not to melodic attention.

Regardless of why we see preferential activation in the semantic attention conditions but not the melodic attention conditions, there is obviously differential processing of the same stimuli when attention is directed toward different aspects of the stimuli, suggesting separate processing of lyrics and melody in song. Attending to semantics involves the LIF area, while attending to melodies does not.

Subtractions specifically looking at the effect of attention support this interpretation. When comparing the two conditions in which stimuli were doubly congruous, different patterns of activation were seen when attending to the semantics than when attending to the melodies. The LIF was more active when listeners attended to the semantics, in addition to a right orbital frontal area. The LIF can be said to be specific to the semantic processing because it is never seen in the melodic attention scans. We cannot say the same about the orbital frontal area, as it is seen in other subtractions including some melodic attention scans. Looking at the flipside of this comparison, when attending to the melodies in the doubly congruous condition, two areas were more active than when listeners attended to the semantics; left superior parietal lobe, and right posterior cingulate. This cingulate area was never significantly more active in semantic conditions than melodic conditions. However, since the functional significance of cingulate activity is currently poorly understood, and the cingulate appears to subserve a variety of different functions (Paus et al., 1998; Peterson et al., 1999), we are unable to judge the importance of this finding at this time.

When we averaged all three of the semantic attention scans and compared them to an average of the three melodic scans, thereby increasing statistical power, some of the same areas emerged, in addition to others. The LIF, along with the right orbital frontal area, was again more active when attending to semantic differences, while the right posterior cingulate was more active when attending melodic differences, replicating the activations seen with the single scan comparison described above. In addition, semantic attention produced an activation in the left posterior temporal lobe, while melodic attention activated a right posterior temporal region (See Figure 2-B).

This left posterior temporal region appears to be part of a network of regions involved in language processing. It has been noted by other imaging studies presenting words in either the visual or auditory modality, in both phonological (Petersen et al., 1988; Petersen et al., 1989; Zatorre et al., 1992) and semantic (Vandenberghe et al., 1996; Binder et al., 1997) tasks. For example, one study activated this region by comparing a condition in which subjects determined the category of three words / pictures (semantic task) to a control condition in which they judged size of the font / pictures (Vandenberghe et al., 1996). So it makes sense that we should see this area, apparently related to language processing, more involved when our listeners attended to the lyrics of the stimuli as opposed to their melodies.

The posterior right temporal activation seen to be more active during melodic monitoring is consistent with other studies in that musical stimuli tend to generate right temporal lobe activity. It is interesting that the area emerges even when the comparison task involves listening to the exact same stimuli, indicating that this area may be related to the directing of attention toward or extraction of melodic aspects of sounds. Indeed, a similar area in posterior middle temporal gyrus has been described as belonging to a network of right hemisphere cortical regions utilized when attending to either tonal frequency or spatial location (Zatorre et al., 1999). The particular focus seen in our comparison is more posterior than is usually observed, nearing the temporo-parieto-occipital junction (see also Binder et al, 1997). This difference is probably due to differences in experimental tasks used in other studies, as well as differences in comparison tasks. For example, in our study, subjects monitored complete musical phrases, while many experimental musical tasks involve comparing isolated tones or short series of tones. Although this particular region of the temporal lobe has not been seen very often, it is not unusual to see temporal activation outside primary and secondary auditory cortical areas when the task involves more complex demands than simple frequency discrimination. For instance, a PET study in which musicians decided whether two successive tones formed minor or major thirds elicited bilateral middle/inferior temporal activity when compared to a noise baseline (Zatorre et al., 1998). Future research will allow us to examine the functional role of the posterior middle temporal region.

Another study specifically comparing an auditory semantic condition to a tonal condition, as in the present study, found temporal activations similar to ours (Binder et al., 1997). The semantic task, determining if auditorily presented animal names were both native to the US and used by humans, was compared to a tonal task in which listeners determined if two high tones were present in short tone series. Although these tasks differ from the congruity decisions our subjects made, comparing their semantic and tonal conditions revealed two posterior temporal regions similar to those in our comparison: the left posterior STG was more active in the semantic condition, and the right posterior middle temporal gyrus was more active in the tonal condition. These posterior temporal regions of activity appear to overlap with the regions seen in our study, and support the idea that they are involved in the directing of attention toward either semantic or tonal aspects of an auditory stimulus.

Congruity Effects:

Another goal of this study was to try to illuminate the structures implicated by evoked potential studies that are differentially engaged when incongruities are present vs. when the stimuli are congruous (Besson & Faita, 1995; Janata, 1995; Besson et al., 1998). Specifically, we were interested in investigating which structures contribute to the N400 seen with semantic incongruities, and those that contribute to the LPC seen with melodic incongruities. Although we used the same stimuli that elicited these responses with evoked potentials, our analyses were unsuccessful in illuminating the neural generators of these responses.

Perhaps this non-result is due in part to the nature of our paradigm. The PET methodology necessitates averaging over a full minute of activation. Due to the long durations of the musical stimuli, we were not able to include many excerpts within each scan. Since there was only one critical event at the end of each excerpt, most of each scan was spent simply listening to the non-critical context building up to the critical events. The time spent processing the critical events was quite short in comparison to the amount of time spent listening to the excerpts. With the EEG evoked potential methodology, it is possible to isolate neural responses to each critical event, cutting out responses to insignificant events, such as the response to all but the last event of the musical excerpt. Also, this evoked potential methodology involves averaging over many trials. We were only able to present each condition once to each subject, with each condition containing only six critical events, limiting our ability to measure responses to the events of interest.

It is also possible that the non-result we see when looking for congruity effects is due to a difference in the type of information obtained with PET as compared to EEG. The EEG samples brain activity with fine timing information but its spatial accuracy is not very exact, while PET uses a very rough time-scale, but is much more precise when it comes to location of activation. The N400 response seen in Besson et al's data using the same stimuli lasted for approximately 550 msec, while the LPC lasted around 800 msec (1998). In our incongruous scans, only six incongruous stimuli could be presented over the course of our one minute of scanning. So given the duration of neural responses to the incongruous stimuli, as measured by EEG, the involved structures would only be responding in this way between 5-8% of each one-minute scanning period. Structures involved in attending to the melody or semantics of the stimuli, on the other hand, would be active whenever the stimuli were being heard, or around 83% of each scan, creating a signal to noise ratio in which we would be more likely to extract the signal. Perhaps the question of congruity effects could be addressed more productively with a different methodology, such as auditory event-related fMRI (e.g. Belin et al., 1999).

Conclusion

In conclusion, this study argues for separate processing of lyric and melody in song. The same song stimuli produced markedly different patterns of results when attending to the lyrics as compared to the melodies. In keeping with the language literature, the LIF region was active when listeners made semantic judgments, while this area was never active during melodic judgment conditions. In addition, the left posterior STG, another known language area, was more active when monitoring semantic than melodic incongruities, while the right posterior middle temporal gyrus was more active when monitoring melodic incongruities, in keeping with studies finding right temporal cortex more involved in music processing.

The finding that lyrics and melody are processed by separate neural systems is consonant with Besson et al's previous study using EEG to examine the issue (1998), along with other studies described above. However, it conflicts with Serafine et al's series of behavioral studies (1984) which consistently find lyrics and melody to be integrated in memory. Perhaps, as has been previously argued, the nature of Serafine et al's task encourages a certain amount of integration in order to facilitate performance on the forced-choice recognition task that would not occur in a natural setting (Peretz et al., 1994). While we do not claim that our study explains all the discrepancies in the literature, we do feel that it makes a strong contribution to the evidence pointing toward separate processing of lyrics and melodies, two intertwining dimensions of song.

Methods

Stimuli: The stimuli used in this study were taken from a collection of sung excerpts of French opera used previously in Besson et al's study (1998). All excerpts were sung a capella (meaning without instruments) by a professional singer. Each was sung four times varying the last event in this manner; the last word/note was either 1) Both semantically and melodically congruous (correct), 2) Semantically congruous and melodically incongruous, 3) Semantically incongruous and melodically congruous, or 4) Both semantically and melodically incongruous. Figure 1 shows a sample excerpt with the four possible endings. Congruous and incongruous words were monosyllabic and matched for frequency of occurrence and length. Whenever possible, incongruous words rhymed with the expected endings. We chose the shortest excerpts for this study (average duration 9-10 seconds) in order to maximize the number of stimuli presented during each scan. All incongruities contained in the chosen excerpts were found to be highly detectable in a behavioral study (Bonnel et al., Submitted).

We normalized the selected excerpts for intensity before re-recording them back onto DAT with a 16 bit, 20,000 Hz sample rate. Interstimulus intervals were two seconds in duration. During this process we created four parallel stimulus sets. Within each set, no excerpt was repeated in any form.

Participants: All 11 participants (4 male, 7 female) were right-handed francophone musicians ranging in age from 22 to 40 All of them had previously been exposed to the stimuli by participating in a behavioral study using the same stimuli (Bonnel et al., Submitted). All subjects gave informed consent and were compensated monetarily for their time and attention.

Procedure: Each participant was randomly assigned to one of four equivalent stimulus sets. For each of six test scans, they were instructed to attend either to the lyrics or to the melodies, and decide if the last event of each excerpt was congruous or incongruous in the relevant dimension (See Table 1 for a list of conditions). Each participant was scanned once during each of these six conditions, in addition to a noise baseline scan in which they listened to white noise fit to the amplitude envelope of some of the stimuli.

During each scan, seven excerpts were presented with one foil in the 3rd, 4th or 5th position. This set of seven was always flanked by two more foils, one each before and after the scan, so that for each block of trials, subjects heard nine excerpts, seven of which were heard during the scan, with one of these as a foil. We included foils to ensure the subjects were paying attention and not simply responding mechanically. Participants heard three practice trials before each scan. They responded via mouse press indicating if each excerpt was congruous or incongruous on the relevant dimension, and responded with random mouse presses after each presentation of noise in the baseline condition.

Imaging and analysis: Sixty-second PET scans were obtained with a Siemens Exact HR+ scanner operating in 3-D mode measuring cerebral blood flow (CBF) using the standard water bolus technique (Raichle et al., 1983). PET images were reconstructed using a 14 mm Hanning filter. These data were then normalized for global CBF value (Evans et al., 1991). T-statistic volumes were computed by dividing each voxel by the pooled standard deviation in normalized CBF for all intracerebral voxels (Worsley et al., 1992). All subjects also received an anatomical T1 weighted MRI scan for co registration with the PET scans, enabling precise anatomical localization. All scans were resampled into standardized stereotaxic space (Talairach & Tournoux, 1988). Scans for each condition were averaged across subjects to increase the signal to noise ratio. Differences between conditions were tested using the subtraction method. Most comparisons were made between two single group-averaged scans. To further increase statistical power, we also created an average of all three group-averaged semantic-attention scans and of all three group-averaged melodic attention scans, and compared the resulting data using the subtraction method. The significance of focal changes in CBF was tested by a method based on 3-D Gaussian random-field theory (Worsley et al., 1992). Unpredicted results were deemed significant at t >= 3.53, based on a whole-brain search region. Predicted areas were deemed significant at lower t-values based on the size of the region of interest.

Chapter 6

General Discussion

Summary and integration of the current findings

Conclusions from behavioral study on normal population

Results from the behavioral study with normal listeners verified the existence of a contextual effect on pitch perception in that the tonal context in which a tone is heard may change the way the pitch of that tone is perceived. In this study, a pitch constancy task was presented in which listeners determined if the pitch of two tones, perhaps differing in spectral shape, was the same or different. The task, although nearly impossible when presented in isolation, was made much easier by placing it within a tonal context. In isolation, spectral shape dominated pitch judgments, while melodic context allowed listeners to better focus in on the frequency dimension and extract differences in F0.

A progressive improvement in pitch discrimination was seen with three increasing levels of context. The extra tones heard in the repeating and alternating tone series conditions, as compared to the isolated condition, created a stronger reference point from which to judge the test tone, thus enhancing pitch discrimination. Expectation of the final tone, created by the repetitive nature of the series, was also a facilitating factor. However, the melodic condition produced the best pitch discrimination. The facilitation seen in this condition as compared to the tone series condition appeared to be due to the structured tonality of the melodies. In addition to the expectation created by predictable melodic progressions, the rich internal structure of these melodies, based on the rules of the Western tonal system, offers more tonal information to the listener than simple repeating or alternating sequences. The auditory system takes advantage of this extra information when extracting pitch both with and without conflicting spectral information. In establishing the effect of tonal context on pitch perception, and operationalizing a way of studying it, this study created a foundation for the following two studies. It also served to develop a paradigm suitable for testing a patient population.

Conclusions from behavioral study on patient population

The next study in this series used essentially the same paradigm to test patients with focal brain excisions in right or left anterior temporal lobe on pitch discrimination in isolated and melodic contexts. As in the behavioral study summarized above, all subjects had difficulty performing the pitch extraction task in the isolated condition. Placing the test tones within the context of a melody improved performance for all three groups (RT, LT, and NC). However, the degree of improvement varied between groups. In particular, patients with left temporal lobe damage performed as controls, while patients with corresponding lesions in the right hemisphere were not able to obtain the same amount of facilitation from the melodic context. We suggested two factors that could be contributing to the impairment seen in the RT group when discriminating pitch in the melodic condition; 1) difficulty holding the contextual tones in memory while waiting for the test tone, and 2) impaired processing of the melodic information. Similar patients to those in the RT group have previously exhibited impairments when holding tones in working memory over a short time interval (Samson & Zatorre, 1991; Zatorre & Samson, 1991). We therefore proposed that the RT group may be having difficulty holding the contextual tones in memory long enough for them to affect perception of the test tones, and are thus not obtaining the full benefit of melodic context. Several other studies implicating right auditory cortical regions in processing melodic contour (Zatorre, 1985; Samson & Zatorre, 1994; Liégeois-Chauvel et al., 1998) led us to suggest that impaired melodic processing by the RT group may prevent them from obtaining the same facilitating information from the melodies as the other groups.

Performance in the RT group did not differ if the lesion included or excluded Heschl's gyrus. We had predicted that patients whose right temporal excision included this region would be more impaired when the discrimination required extracting pitch from tones differing in spectral shape than when this process was not necessary. Looking at results from the isolated condition only to avoid any confounding effects of context, no effect was seen on same- or different-timbre trials when excisions extended into Heschl's gyrus or not. However, this negative result may be due to a floor effect seen across groups in the isolated condition in which all groups, including controls, performed near chance. So although no effect of spectral shape was seen in this study, it is unclear whether this lack of impairment is due to spared auditory processing, or to a floor effect seen across groups that may make it difficult to detect an impairment.

Since no dissociation of responses was found between RT patients whose excision included or spared Heschl's gyrus, the deficits seen in the RT group were attributed to the excision of more anterior areas of the temporal lobe, excised in all RT patients. This study therefore implicated anterior portions of the auditory cortex in making pitch judgments relative to tones that were heard previously. From these results, we proposed that auditory association areas located on the anterior portion of the STG, which have connections to dorsolateral frontal regions implicated in working memory tasks (Chavis & Pandya, 1976; Pandya & Yeterian, 1985; Petrides & Pandya, 1999; Romanski et al., 1999a), could be involved in holding contextual tones of melodic context in memory while listeners wait for a final test tone.

Adapting behavioral paradigm for fMRI

The patient study generated some interesting findings that merited deeper investigation. In particular, the conclusion that the right anterior STG is able to process tones relative to tones that were heard previously was very intriguing. In order to verify and extend these results to a normal musically inclined population, the paradigm was adapted for use with fMRI. As the object of investigation was to determine neural substrates involved in creating context effects on pitch perception, the spectral difference aspect of the paradigm was dropped, and all tones were presented with the same spectral timbre. As well, a random tone series context was included as an intermediate condition between the isolated and melodic contexts.

Differences between the way behavioral responses were measured in this paradigm and those used in the previous two studies deserve mention. For instance, the basic analysis measures, previously accuracy and pitch difference ratings, were replaced with reaction time. This was changed due to a shift in how task difficulty was accounted for. Since previous functional neuroimaging studies have shown increases in neural activity due solely to increases in task difficulty (e.g. Winstein et al., 1997), we attempted to control for this potential confound by keeping difficulty level constant across subjects. This was achieved by measuring individual pitch discrimination thresholds, and presenting different-pitch trials at a constant difference above threshold. This level was chosen so that pitch differences would be relatively obvious, thus reducing incorrect trials. However, this manipulation makes accuracy measures uninformative, necessitating the use of reaction time.

Conclusions from fMRI study with musician population

Basing our hypothesis on the previous patient study, we predicted increasing activity in the auditory association areas located on the right anterior STG as a function of increasing context. We also predicted right dorsolateral frontal regions to behave similarly, due to its previous implication in tonal working memory tasks, and its anatomical connections to association areas on the anterior STG. As predicted, within the extensive bilateral STG activity evident in each contextual condition compared to baseline, right anterior STG activation increased with increasing tonal context. Right dorsolateral frontal activation, also present in each contextual condition, remained at a constant level across contexts, indicating a similar amount of processing in this region whenever pitch discriminations are required, regardless of contextual information.

In this study, as in the previous two, melodic context produced the most facilitation of pitch discrimination. This result, although predicted, was welcomed in that it provided converging evidence of context effects on pitch perception with a different behavioral method. Whereas the previous two studies used accuracy and pitch difference ratings to show this effect, the fMRI study used reaction time while listeners performed at ceiling accuracy levels. However, the random condition, meant to be an intermediate context between the other two, was not performed significantly faster than the isolated condition. The prediction was that as in the repeating and alternating tone series conditions in the behavioral study, random series of tones would produce results that fell between the isolated and melodic conditions. Although pilot testing indicated the predicted pattern of results, the final analysis did not even detect a trend in this direction. Random tones taken from the chromatic scale as they were in this study, although unstructured, reduce the set of possible discrete F0's to only those contained in the chromatic scale, thus offering more information about the test tones than the isolated condition. However, unlike the melodies and alternating/repeating tone series which do show facilitation of pitch discrimination, random tones do not offer any information useful for predicting the pitch of the test tones. Therefore, the lack of pitch facilitation in a random tonal context suggests that the expectation of particular pitches set up by melodies and predictable tone series is an important facilitating factor of these contexts.

Combining the functional and behavioral results of the fMRI study allows a more comprehensive interpretation than either result alone. Melodic context produced the most facilitation behaviorally, as well as the most activity in right anterior STG. This pattern of behavioral and functional results fits well with the theory that auditory processing in the right anterior STG contributes to the effect of tonal context on pitch perception. However, whereas the random context produced activation levels intermediate to the isolated and melodic conditions in the same area, this increased activity was not translated into behavioral facilitation.

This apparent contradiction can be partially explained by two related points. First, the results of this study do suggest that the right anterior STG is more active with increasing tonal information. As previously described, the random tone series contains more tonal information than the isolated condition, and less than the melodic condition. Just because this information does not lead to a behavioral facilitation of pitch judgment does not mean it is not being processed by the auditory system in some manner. The intermediate activity in right anterior STG during the random condition relative to isolated and melodic conditions suggests that this information is being processed somehow, but the current behavioral paradigm is not sensitive to any perceptual changes produced by this neural activity. In a related point, we have merely observed a correlation between the behavioral and functional results, and cannot deduce that changes in anterior STG activity cause the behavioral differences seen between conditions. Although they appear to be related, the exact nature of this relation is unknown, and cannot be inferred without further study. Other factors, currently unknown, are certainly involved in this interaction between neural activity and the resulting perception and response.

The main finding of this study was that auditory association areas on the right anterior STG appear to have a role in incorporating contextual information into pitch perception. This fits in well with data from the lesion study finding patients with resections of this area to have impaired use of contextual cues in a similar tone discrimination task. This region's connections to and co-activation with right dorsolateral frontal cortex, a region involved in tonal working memory, appear to allow processing of tones relative to those heard previously. Activation of right anterior STG increased significantly with each level of increased tonal context, with melodic context generating the most activity, and producing the quickest pitch discrimination behaviorally, suggesting that neural processing in this region is related to a facilitation in pitch discrimination.

Conclusions from PET study with musician population

While addressing some similar issues to those addressed in the previous three studies, the PET study described in Chapter 5 does not follow directly from them. This study used a question currently being debated in the literature to broaden the scope of investigation to include linguistic aspects of audition. The question was whether or not the lyrics and melody of song are processed separately within the auditory system. A task similar to the melodic condition of the previous three studies was employed, with half of the conditions requiring listeners to judge if the final note of an opera excerpt was sung at the correct pitch, and the other half requiring judgment of whether the last word of the excerpt was semantically correct. So instead of comparing a melodic condition to another tonal condition, the melodic conditions were compared to linguistic conditions. This paradigm enabled direct comparison of melodic and semantic processing of the same stimulus.

Due to the similar stimuli used in all conditions, sung excerpts of opera, similarities in the pattern of activation were seen across conditions. For example, extensive bilateral STG activity was seen in all test minus baseline comparisons, a result also seen in the fMRI study described in Chapter 4. This similarity across conditions, as well as across studies, is probably due to the large amount of auditory processing necessary to perform the tasks. However, even though the stimuli were the same, activation differences were also observed between conditions depending on which aspect of the stimulus listeners were required to attend to in order to perform the task. For semantic judgments, listeners were required to attend to the lyrics, but the melodic aspect of the stimuli was irrelevant. When attending to the lyrics, a left inferior frontal region was active, an area associated with linguistic processing (e.g. Petersen et al., 1988; Binder et al., 1997). In addition, a region on the left posterior STG was found to be more active when subjects made semantic rather than melodic judgments. When melodic judgments were required, making the semantic content of the excerpts irrelevant, a different posterior temporal region was active: the right middle temporal gyrus. These different regions of activity appear to be related to attending to different aspects of the stimulus: semantic or melodic. We did not expect, and did not see differential activity in the right anterior STG as in the fMRI study because this area appears to be sensitive to differences in tonal context, and the amount of tonal context available in each of the PET conditions was the same.

The right posterior middle temporal region specific to attending to the melodies in this study was not seen in the fMRI study. However, this is not surprising due to the differences in tasks and comparison conditions between the two studies. In the fMRI study, the melodic condition was either compared to other tonal conditions or to a silent baseline, whereas the PET paradigm compared the melodic conditions to semantic conditions. Also, the stimuli used in the PET study were musically much richer than the stimuli used in the fMRI study, perhaps requiring additional auditory processing. But regardless of these task differences, discrepancies in activation patterns across studies should not be interpreted too strictly, as different effects were subtracted out through the use of different comparison conditions.

Even though direct comparisons should not be made between the PET study and the other studies included in this thesis, the PET study, while still requiring auditory judgments to be made with respect to contextual information, adds another dimension in that it addresses differences between tonal and linguistic auditory processing. By focusing specifically on certain tonal aspects of auditory processing, the first three studies reported in this thesis are able to deal with specific features of this processing in detail. By broadening the investigation to include linguistic aspects of auditory processing, the PET study allows the information gained from the previous studies to be put into a more global perspective. It is then possible to see how the tonal processing fits into the wider domain of general auditory processing. In particular, the PET study showed that although the auditory processing of tonal and linguistic aspects of sounds share some common neural substrates, some of these substrates are used more or less in each type of processing, such as the two posterior temporal regions, and some are only recruited for one process or the other, such as the left inferior frontal region. So although it does not deal specifically with the effects of tonal context on pitch perception, the PET study adds to the other studies by broadening their scope and placing them in a wider auditory perspective.

Integration of functional findings with previous literature

Both the fMRI and PET studies reported in this thesis showed extensive activity throughout the STG bilaterally when comparing task conditions to baseline, a result also seen in other studies using complex auditory stimuli (e.g. Zatorre et al., 1996). These gyri contain the human auditory system's primary, secondary, and association cortices, and comprise of a variety of regions distinguishable by differences in cytoarchitecture (Galaburda & Sanides, 1980), connectivity (Pandya & Yeterian, 1985), and response properties (Liégeois-Chauvel et al., 1994). In addition, damage to different regions of the STG produce different auditory impairments (e.g. Samson & Zatorre, 1991), and electrophysiological studies in animals such as the monkey and the cat have shown neurons in different auditory regions to exhibit different neural coding patterns for various sounds (Rauschecker et al., 1995; See Brugge & Reale, 1985 for review). Therefore, the broad activation of bilateral STG in the fMRI and PET studies includes multiple distinct auditory regions, and is consistent with the processing of many various aspects of the complex stimuli used in these studies.

Temporal lobe regions outside the STG (Zatorre et al., 1994; Binder et al., 1997; Belin et al., 2000) have also been shown to respond to auditory stimuli, such as the superior temporal sulcus, and the middle temporal gyrus. Unlike primary and secondary auditory cortices, these regions may not respond exclusively to auditory stimuli (Grafton et al., 1996), and appear to be involved in processing more complex auditory information, such as voices (Belin et al., 2000) or counting occurrences of a specific tone (Binder et al., 1997). The right posterior middle temporal gyrus activity seen in the current PET study to be more active when attending to melodies than semantics emphasizes the point that not all auditory processing is performed in the STG, and that higher order auditory perception may be influenced by processing in other temporal-lobe regions.

Right anterior STG and Right dorsolateral frontal regions

Both the behavioral study of patients and the fMRI study of musicians suggest an involvement of the auditory association area located on the right anterior STG in producing effects of tonal context on pitch perception. In keeping with the idea that association areas are more involved in higher-order stimulus processing than primary and secondary cortices, this region appears to be involved in the complex task of perceiving tones relative to tones that were heard previously. To my knowledge, this is the first time anyone has suggested a specific region in the human auditory cortex capable of processing tones in such a way.

This anterior temporal region appears to work in tandem with ventrolateral, or opercular frontal regions in both hemispheres, as well as right dorsolateral frontal regions specifically; each of the three contextual conditions for pitch discrimination in the fMRI study showed consistent right dorsolateral frontal activity. In this study, the right dorsolateral and anterior STG regions responded differently to the differing tonal contexts: increasing tonal context caused right anterior STG activity to increase, whereas right dorsolateral frontal activity stayed constant across contexts. This right frontal activity was interpreted as being involved in tonal working memory. However, if this region was holding specific tones in memory, more activity should have been seen there in the melodic and random conditions than in the isolated condition, due to the need to keep more tones in memory. Since this did not occur, and a consistent level of activity was seen in all three contexts, it is possible to speculate that perhaps the tonal context is coded in some way in neural processes carried out in the right anterior STG, and that this coded information is then processed in working memory via interactions between these two areas.

Cortico-cortical connections have been traced between this frontal region and the anterior STG in monkeys (Petrides & Pandya, 1988; Petrides & Pandya, 1999; Romanski et al., 1999a; Romanski et al., 1999b). Cytoarchitectonic studies outlining analogous cortical regions between human and non-human primates have been very helpful in allowing the extensive anatomical work performed with monkeys to be applied to human anatomy, of which far fewer studies exist (Galaburda & Sanides, 1980; Petrides & Pandya, 1994; Petrides & Pandya, 1999). Due to the involvement of dorsolateral frontal cortex in working memory, the connections seen between the right dorsolateral frontal region and the anterior STG may be a mechanism for holding contextual tones in memory (Perry & Marin, 1999).

One theory of frontal lobe involvement in working memory involves a hierarchy of processing (Petrides, 1994). In this system, simply holding a low number of items in memory does not require frontal-lobe activity. However, making judgments on these items requires ventrolateral frontal structures, and monitoring or manipulating these items requires additional activity in dorsolateral frontal regions, usually in areas 9, 9/46, and 46 (Owen, 1997; Petrides & Pandya, 1999).

The present functional studies in Chapters 3 and 4 lend only partial support to this theory. For example, all three contextual conditions in the fMRI study show ventrolateral frontal activity, consistent with judging the pitch of tones in working memory, but instead of favoring right frontal structures as predicted, this activity is seen either on the left only, as in the melodic condition, or bilaterally. These three conditions also all exhibit strong right dorsolateral frontal activity, with no corresponding activity in left frontal regions, consistent with monitoring the contextual tones. However, in the case of the isolated condition, this dorsolateral frontal activity is present even though there are no contextual tones to monitor. The reason for this result is unclear, but finding activity in right dorsolateral cortical regions when comparing the pitch of two isolated items is not unprecedented (Zatorre et al., 1992; Holcomb et al., 1998).

In the present PET study, listeners monitored melodic and semantic events, and made a judgment on one of these events relative to the others. However, this task differs substantially from other studies showing dorsolateral frontal activity in that no explicit comparison is made between items contained in working memory, and indeed, no dorsolateral frontal activity was observed in this study. Instead, congruity judgments are made regarding how one of these items follows from the others. The working memory theory could therefore predict that making this judgment relative to the contextual events held in working memory would recruit ventrolateral, or inferior, frontal regions. Indeed, the LIF region was active in all conditions requiring semantic judgments, but no corresponding activity was seen in either frontal lobe when melodic judgments were required.

Despite the theoretical complexities regarding the exact nature of differences between processing within dorsolateral and ventrolateral frontal cortex, these regions do appear to be involved in tonal working memory in some manner. Right frontal regions in particular have been implicated in many different tonal working memory tasks (Zatorre et al., 1992; Zatorre et al., 1994; Zatorre et al., 1996; Binder et al., 1997; Holcomb et al., 1998; Zatorre et al., 1999; see Perry and Marin, 1999 for a review). This common result is evident even though the paradigms used in these studies were all slightly different from each other. For example, in one study, listeners determined if the pitch of two spoken syllables rose or fell (Zatorre et al., 1992). Another study required responses to occurrences of a particular high tone only, where the frequency difference between the critical and noncritical tones was large (Zatorre et al., 1999). Still another study required trained listeners to label each tone in a series as having high or low pitch when the difference between these tones was close to threshold (Holcomb et al., 1998). The convergence of results across studies and the present studies strongly supports the idea that these right frontal regions are involved in tonal working memory.

When pitch discrimination requires comparing tones that are separated by a series of other tones, thus increasing the tonal memory load, right dorsolateral frontal activations are joined by similar activations in the left frontal lobe (Zatorre et al., 1994; Zatorre et al., 1996). In both of these studies, however, a rightward asymmetry of frontal activity is seen in that more activity is evident on the right than on the left. This bilateral frontal activity seen with increased memory load may be due to the increased difficulty of the task. The authors of one of these studies suggests that the additional frontal activity produced by these more difficult tasks may be related to mental processes other than working memory that are required by the task, such as monitoring the tone presentation along with the temporal order of the tones, a process not necessary when simply comparing the pitch of two isolated tones (Zatorre et al., 1994).

The similarity of tasks in one of these studies compared to the conditions used in this thesis, makes it an interesting subject for comparison (Zatorre et al., 1994). In this

PET study, one task involved comparing the pitch of the first and last tones of short 8-tone melodies (first/last). This task was similar to the present melodic conditions in that a pitch judgment was made within the context of a melody. The second task required a pitch comparison of the first two tones of these melodies, in which a decision was made before the tonal context was presented (2-tone), thus making this condition similar to the present isolated condition. However, instead of judging small pitch differences as in the present studies, this PET study used a slightly different task requiring listeners to determine if large pitch differences (average 5 semitones, or 500 cents) increased or decreased.

The 2-tone judgment, when compared to passive listening of the melodies, generated increased activity in two right frontal lobe regions, with no corresponding frontal activity in the left hemisphere. When the task was made more difficult by requiring listeners to hold the first tone in memory during the entire melody presentation (first/last condition), two dorsolateral frontal regions similar to those seen in the fMRI study were preferentially active on the right, as well one similar region on the left. The left frontal region that was recruited in addition to the right frontal regions could be related to the increased memory component of this first/last condition. Alternatively, as mentioned above, this left frontal activity could be associated with additional mental processes required when monitoring a sequence of tones. The authors of this study suggested that the higher memory load recruited a specialized auditory working memory system that includes interactions between right frontal and temporal regions, a similar conclusion to the one put forth in this thesis.

Rightward asymmetry of neuronal activation during pitch processing

In accordance with previous studies (Milner, 1962; Zatorre, 1985; Zatorre, 1988; Divenyi & Robinson, 1989; Sidtis et al., 1989; Robin et al., 1990; Zatorre & Samson, 1991; Zatorre et al., 1992; Zatorre & Halpern, 1993; Zatorre et al., 1994; Zatorre et al., 1996; Binder et al., 1997; Holcomb et al., 1998; Zatorre et al., 1999; Johnsrude et al., 2000), all three experiments examining the neural correlates of pitch perception in this thesis showed a rightward asymmetry of neuronal processing in the temporal and frontal lobes. This result is consistently found in the literature across widely varying paradigms and measuring techniques. However, pitch processing does not appear to be carried out solely in right temporal and frontal regions. Impairments on pitch-related tasks are sometimes seen after damage to left auditory structures, but damage to right auditory regions nearly always causes a larger detriment to pitch processing (Zatorre, 1985; Peretz, 1990; Samson & Zatorre, 1991; Liégeois-Chauvel et al., 1998). In addition, bilateral lesions of auditory structures appear to result in the greatest deficits on pitch-related tasks (Peretz et al., 1994). Therefore, the rightward asymmetry of pitch processing appears to be a matter of degree, and not strict laterality *per se*.

Conclusion

In sum, the combination of the four studies reported here provides a comprehensive contribution to the literature examining the neural correlates underlying pitch perception in general, and how tonal context affects this percept specifically. After establishing the perceptual effect of tonal context on pitch through behavioral testing of a normal population, the cortical structures giving rise to this effect were investigated by testing patients with focal unilateral temporal lobe excisions behaviorally, as well as testing musicians with a functional neuroimaging methodology used in conjunction with behavioral methods. These studies found converging evidence that a network of structures involving the right anterior STG and right dorsolateral frontal regions contributes to these contextual effects.

The neural correlates of pitch processing were then contrasted with those of semantic processing in another functional neuroimaging study that compared neuronal activity due to melodic and semantic judgments of song stimuli. Semantic processing was found to involve increased activation in left temporal and frontal lobes, while the importance of right temporal regions in the processing of pitch was again emphasized.

The integration of these four studies highlights the benefits of finding converging evidence through multiple methodologies. A deeper understanding of the neural systems subserving complex perceptions can be gained by exploiting the unique strengths of different neuropsychological techniques. These studies suggest a variety of further research projects, including the integration of timing information obtainable with ERP methodologies with the functional data reported here. This type of study could elucidate the temporal progression of activation of structures implicated in the present studies. In addition, an ERP study could provide information regarding the timecourse of contextual effects on the perception of a tone. Another possible extension of the work in this thesis involves examining the role of right frontal regions in producing contextual effects on pitch by testing patients with focal unilateral excisions in the frontal lobe.

Clarification of the precise nature of the anterior STG in processing auditory events relative to contextual information could be addressed by contrasting the present tonal contextual paradigm with analogous linguistic contextual conditions. One possible suggestion would be to require listeners to discriminate specific phonemes presented in varying levels of linguistic context. This type of study could clarify whether the right anterior STG region seen increasing levels of tonal context is located on the right because it is processing tonal information, and a similar region would be active on the left if linguistic context was varied, or if this region would be seen on the right with increasing contextual non-tonal information as well.

The present studies underscore the value of using tonal stimuli to elucidate the complexities of auditory processing. By combining behavioral and functional studies examining tonal or musical processing with those examining that of other auditory stimuli such as linguistic or environmental sounds, it will be possible to attain a more complete understanding of how the auditory system generates the complex perceptions we ultimately experience.

Finally, the perceptual concepts examined in this thesis highlight the idea that perception as we experience it is generated by more than the physical sensory information reaching our nervous system. This physical information is processed in a variety of complex ways, including relating bits of information to each other, which maximizes the final percept obtained by the system. Since the information we receive from our environment is frequently incomplete or degraded, this top-down processing is extremely useful in our day-to-day perceptions. The ability to interpret similar events differently in different conditions, as well as recognize the same object as being the same in differing conditions, and knowing which of these mechanisms to employ at any particular time, is very important for a clear understanding of the environment, but is also a very difficult problem for the perceptual system to solve. In order to make any sense of these problems, it is necessary to be able to interpret contextual cues. In the auditory modality in particular, this interpretation must be performed in the temporal domain; unlike visual objects, auditory events disappear after they occur, and cannot be perused at length. This temporal element of audition makes the utilization of auditory working memory mechanisms particularly important when interpreting sounds in relation to each other. This thesis offers one possible mechanism for interpreting auditory information in relation to the context in which it is heard.

References

Achim, A., Ahad, P., & Bregman, A. (1992). MAPLE: McGill Auditory Perception and Linguistic Experiments. Montreal, QC: McGill Speech and Hearing Lab.

American National Standards Institute. (1960). *USA standard acoustical terminology.* New York: American National Standards Institute.

Baker, S. C., Rogers, R. D., Owen, A. M., Frith, C. D., J, D. R., Frackowiak, R. S., & Robbins, T. W. (1996). Neural systems engaged by planning: A PET study of the Tower of London task. *Neuropsychologia, 34*(6), 515-526.

Bartholomeus, B. (1974). Effects of task requirements on ear superiority for sung speech. *Cortex, 10*, 215-223.

Belin, P., Zatorre, R. J., Hoge, R., Evans, A. C., & Pike, B. (1999). Event-related fMRI of the auditory cortex. *NeuroImage, 10*, 417-429.

Belin, P., Zatorre, R. J., Lafaille, P., Ahad, P., & Pike, B. (2000). Voice-selective areas in human auditory cortex. *Nature, 403*(20 Jan), 309-312.

Belliveau, J. W., Kennedy, D. N., McKinstry, R. C., Buchbinder, B. R., Weisskoff, R. M., Cohen, M. S., Vevea, J. M., Brady, T. J., & Rosen, B. R. (1991). Functional mapping of the human visual cortex by magnetic resonance imaging. *Science, 254*, 716-719.

Besson, M., & Faita, F. (1995). An event-related potential (ERP) study of musical expectancy: Comparison of musicians with nonmusicians. *Journal of Experimental Psychology: Human Perception and Performance, 21*, 1278-1296.

Besson, M., Faita, F., Peretz, I., Bonnel, A. M., & Requin, J. (1998). Singing in the brain: Independence of lyrics and tunes. *Psychological Science, 9*(6), 494-498.

Bharucha, J. J. (1994). Tonality and expectation. In R. Aiello, and John Sloboda (Ed.), *Musical Perceptions* (pp. 213-239). New York: Oxford University Press.

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

Bigand, E. (1993). Contributions of music to research on human auditory cognition. In S. McAdams & E. Bigand (Eds.), *Thinking in sound: the cognitive psychology of human audition* (pp. 231-277). Oxford: Clarendon Press.

Binder, J. R., Frost, J. A., Hammeke, T. A., Cox, R. W., Rao, S. M., & Prieto, T. (1997). Human brain language areas identified by functional magnetic resonance imaging. *The Journal of Neuroscience, 17*(1), 353-362.

Binder, J. R., Rao, S. M., Hammeke, T. A., Frost, F. A., Bandettini, P. A., Jesmanowicz, A., & Hyde, J. S. (1995). Lateralized human brain language systems demonstrated by task subtraction functional magnetic resonance imaging. *Archives of Neurology, 52*, 593-601.

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

Bonnel, A.-M., Faita, F., Peretz, I., & Besson, M. (Submitted). Divided attention between lyrics and tunes of operatic songs. Evidence for independent processing. *Perception and Psychophysics*.

Bookheimer, S. Y., Zeffiro, T. A., Blaxton, T. A., Gaillard, W. D., Malow, B., & Theodore, W. H. (1998). Regional cerebral blood flow during auditory responsive naming: Evidence for cross-modality neural activation. *NeuroReport, 9*, 2409-2413.

Brugge, J. F., & Reale, R. A. (1985). Auditory Cortex. In A. Peters & E. G. Jones (Eds.), *Cerebral Cortex* (Vol. 4, pp. 229-271). New York: Plenum Press.

Buckner, R. L., Bandettini, P. A., O'Craven, K. M., Savoy, R. L., Petersen, S. E., Raichle, M. E., & Rosen, B. R. (1996). Detection of cortical activation during averaged single trials of a cognitive task using functional magnetic resonance imaging. *Proceedings of the National Academy of Sciences, USA, 93*(Dec), 14878-14883.

Burns, E. M. (1999). Intervals, scales, and tuning. In D. Deutsch (Ed.), *The Psychology of Music* (2 ed., pp. 215-264). San Diego: Academic Press.

Carpenter, P. A., Just, M. A., Keller, T. A., Eddy, W. F., & Thulborn, K. R. (1999). Time course of fMRI-activation in language and spatial networks during sentence comprehension. *NeuroImage, 10*(2), 216-224.

Catalan, M. J., Ishii, K., Honda, M., Samii, A., & Hallett, M. (1999). A PET study of sequential finger movements of varying length in patients with Parkinson's disease. *Brain, 122*(3), 483-495.

Chao, L. L., & Knight, R. T. (1997). Prefrontal deficits in attention and inhibitory control with aging. *Cerebral Cortex, 7*(1), 63-69.Chavis, D. A., & Pandya, D. N. (1976). Further observations on corticofrontal connections in the rhesus monkey. *Brain Research, 117*, 369-386.

Crowder, R. G. (1993). Auditory Memory. In S. McAdams (Ed.), *Thinking in Sound* (pp. 113-145).

Crowder, R. G., Serafine, M. L., & Repp, B. (1990). Physical interaction and association by contiguity in memory for the words and melodies of songs. *Memory and Cognition, 18*(5), 469-476.

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

Demany, L., & Semal, C. (1993). Pitch versus brightness of timbre: Detecting combined shifts in fundamental and formant frequency. *Music Perception, 11*(1), 1-14.

Demb, J. B., Desmond, J. E., Wagner, A. D., Vaidya, C. J., Glover, G. H., & Gabrieli, J. D. E. (1995). Semantic encoding and retrieval in the left inferior prefrontal cortex: A functional MRI study of task difficulty and process specificity. *The Journal of Neuroscience, 15*(9), 5870-5878.

Deutsch, D. (1970). Tones and numbers: Specificity of interference in immediate memory. *Science, 168*(3939), 1604-1605.

Deutsch, D. (1972a). Effect of repetition of standard and comparison tones on recognition memory for pitch. *Journal of Experimental Psychology, 93*, 156-162.

Deutsch, D. (1972b). Mapping of interactions in the pitch memory store. *Science, 175*(4025), 1020-1022.

Deutsch, D. (1982). The Processing of Pitch Combinations. In D. Deutsch (Ed.), *The Psychology of Music* . New York: Academic Press.

Deutsch, D. (1999). The processing of pitch combinations. In D. Deutsch (Ed.), *The Psychology of Music* (2 ed., pp. 349-411). San Diego: Academic Press.

Deutsch, D., & Roll, P. L. (1974). Error patterns in delayed pitch comparison as a function of relational context. *Journal of Experimental Psychology, 103*(5), 1027-1034.

Dewar, K. M., Cuddy, L. L., & Mewhort, D. (1977). Recognition memory for single tones with and without context. *Journal of Experimental Psychology: Human Learning and Memory, 3*(1), 60-67.

Divenyi, P. L., & Robinson, A. J. (1989). Nonlinguistic auditory capabilities in aphasia. *Brain and Language, 37*, 290-326.

Dowling, W. J. (1999). The development of music perception and cognition. In D. Deutsch (Ed.), *The Psychology of Music* (2 ed., pp. 603-625). San Diego: Academic Press.

Ehresman, D. E., & Wessel, D. L. (1978). *Perception of timbral analogies* . Paris: Centre Georges-Pompidou: IRCAM.

Evans, A. C., Thompson, C. J., Marrett, S., Meyer, E., & Mazza, M. (1991). Performance evaluation of the PC-2048: A new 15-slice encoded-crystal PET scanner for neurological studies. *IEEE Trans Med Imag, 10*, 90-98.

Fox, P. T., Mintun, M. A., Herscovitch, P., & Raichle, M. E. (1984). A noninvasive approach to quantitative functional brain mapping with $H_2{}^{15}O$ and positron emission tomography. *Journal of Cerebral Blood Flow and Metabolism, 4*, 329-333.

Friston, K. J. (1997). Imaging cognitive anatomy. *Trends in Cognitive Sciences, 1*(1), 21-27.

Galaburda, A., & Sanides, F. (1980). Cytoarchitectonic organization of the human auditory cortex. *The Journal of Comparative Neurology, 190*, 597-610.

Garner, W. (1974). *The processing of information and structure*. Potomac, MD: Erlbaum.

Goodglass, H., & Calderon, M. (1977). Parallel processing of verbal and musical stimuli in right and left hemispheres. *Neuropsychologia, 15*, 397-407.

Grafton, S. T., Arbib, M. A., Fadiga, L., & Rizzolatti, G. (1996). Localization of grasp representation in humans by positron emission tomography 2. Observation compared with imagination. *Experimental Brain Research, 112*, 103.

Grey, J. (1977). Multidimensional perceptual scaling of musical timbres. *Journal of the Acoustical Society of America, 61*, 1270-1277.

Griffiths, T. D., Johnsrude, I., Dean, J. L., & Green, G. G. R. (1999). A common neural substrate for the analysis of pitch and duration pattern in segmented sound? *NeuroReport, 10*(18), 3825-3830.

Halpern, A. R., & Zatorre, R. J. (1999). When that tune runs through your head: A PET investigation of auditory imagery for familiar melodies. *Cerebral Cortex, 9*(7), 697-704.

Handel, S. (1993). *Listening: An Introduction to the Perception of Auditory Events.* Cambridge, Massachusetts: The MIT Press.

Henke, W. (1981). MITSYN: A coherent family of command-level utilities for time signal processing. Belmont, MA.

Holcomb, H. H., Medoff, D. R., Caudill, P. J., Zhao, Z., Lahti, A. C., Dannals, R. F., & Tamminga, C. A. (1998). Cerebral blood flow relationships associated with difficult tone recognition task in trained normal volunteers. *Cerebral Cortex, 8*, 534-542.

Janata, P. (1995). ERP measures assay the degree of expectancy violation of harmonic contexts in music. *Journal of Cognitive Neuroscience, 7*, 153-164.

Janata, P., & Reisberg, D. (1988). Response-time measures as a means of exploring tonal hierarchies. *Music Perception, 6*, 161-172.

Johnsrude, I. S., Penhune, V. B., & Zatorre, R. J. (2000). Functional specificity in the right human auditory cortex for perceiving pitch direction. *Brain, 123*.

Klein, D., Milner, B., Zatorre, R. J., Zhao, V., & Nikelski, J. (1999). Cerebral organization in bilinguals: A PET study of Chinese-English verb generation. *NeuroReport, 10*, 2841-2846.

Klein, D., Olivier, A., Milner, B., Zatorre, R. J., Johnsrude, I., Meyer, E., & Evans, A. C. (1997). Obligatory role of the LIFG in synonym generation: evidence from PET and cortical stimulation. *NeuroReport, 8*, 3275-3279.

Krumhansl, C. (1989). Why is musical timbre so hard to understand? In S. a. O. O. Nielzen (Ed.), *Structure and perception of electroacoustic sound and music* (pp. 43-53). Amsterdam: Excerpta Medica.

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

Krumhansl, C. L. (1990). *Cognitive Foundations of Musical Pitch* (Vol. 17). New York, Oxford: Oxford University Press.

Krumhansl, C. L., & Castellano, M. A. (1983). Dynamic processes in music perception. *Memory and Cognition, 11*(4), 325-334.

Krumhansl, C. L., & Iverson, P. (1992). Perceptual interactions between musical pitch and timbre. *Journal of Experimental Psychology: Human Perception and Performance, 18*(No. 3), 739-751.

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

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

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 207*, 203-205.

Lechevalier, B., Platel, H., & Eustache, F. (1995). Neuropsychologie de l'identification musicale. *Rev Neurol, 151*(8-9), 505-510.

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *Journal of the Acoustical Society of America, 49*(2 Suppl 2), 467-477.

Liegeois-Chauvel, C., Musolino, A., Badier, J., Marquis, P., & Chauvel, P. (1994). Evoked potentials recorded from the auditory cortex in man: evaluation and topography of the middle latency components. *Electroencephalography and Clinical Neurophysiology, 92*, 204-214.

Liégeois-Chauvel, C., Peretz, I., Babaï, M., Laguitton, V., & Chauvel, P. (1998). Contribution of different cortical areas in the temporal lobes to music processing. *Brain, 121*, 1853-1867.

Luria, A. R., Tsvetkova, L. S., & Futer, D. S. (1965). Aphasia in a composer. *Journal of Neurological Science, 2*, 288-292.

McAdams, S., Winsberg, S., Donnadieu, S., Soete, G. d., & Krimphoff, J. (1995). Perceptual scaling of synthesized musical timbres: Common dimensions, specificities, and latent subject classes. *Psychological Research, 58*, 177-192.

Melara, R. D., & Marks, L. E. (1990a). Hard and soft interacting dimensions: Differential effects of dual context on classification. *Perception & Psychophysics, 47*(4), 307-325.

Melara, R. D., & Marks, L. E. (1990b). Interaction among auditory dimensions: Timbre, pitch, and loudness. *Perception & Psychophysics, 48*(2), 169-178.

Melara, R. D., & Marks, L. E. (1990c). Perceptual primacy of dimensions: Support for a model of dimensional interaction. *Journal of Experimental Psychology: Human Perception and Performance, 16*(No. 2), 398-414.

Mesulam, M.-M., & Mufson, E. J. (1982a). Insula of the old world monkey. I: Architectonics in the insulo-orbito-temporal component of the paralimbic brain. *The Journal of Comparative Neurology, 212*, 1-22.

Mesulam, M.-M., & Mufson, E. J. (1982b). Insula of the old world monkey. III: Efferent cortical output and comments on function. *The Journal of Comparative Neurology, 212*, 38-52.

Meyer, D. R., & Woolsey, C. N. (1952). Effects of localized cortical destruction on auditory discrimination conditioning in the cat. *Journal of Neurophysiology, 15*, 149-162.

Miller, J. R., and Edward C Carterette. (1975). Perceptual space for musical structures. *Journal of the Acoustical Society of America, 58*(No. 3), 711-720.

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

Moore, B. C. J. (1989). *An Introduction to the Psychology of Hearing* (3 ed.). San Diego: Academic Press.

Moseley, M. E., & Glover, G. H. (1995). Functional MR Imaging. *Neuroimaging Clinics of North America, 5*(2), 161-191.

Owen, A. M. (1997). The functional organization of working memory processes within the human lateral frontal cortex: The contribution of functional neuroimaging. *European Journal of Neuroscience, 9*, 1329-1339.

Pandya, D. N. (1995). Anatomy of the auditory cortex. *Reviews in Neurology, 151*(8-9), 486-494.

Pandya, D. N., & Yeterian, E. H. (1985). Architecture and connections of cortical association areas. In A. Peters & E. G. Jones (Eds.), *Cerebral Cortex* (Vol. 4, pp. 3-61). New York: Plenum Press.

Patel, A., Gibson, E., Ratner, J., Besson, M., & Holcomb, P. (1998). Processing syntactic relations in language and music: An event-related potential study. *Journal of Cognitive Neuroscience, 10*(6), 717-733.

Paus, T., Koski, L., Caramanos, Z., & Westbury, C. (1998). Regional differences in the effects of task difficulty and motor output on blood flow response in the human anterior cingulate cortex: A review of 107 PET activation studies. *NeuroReport, 9*(9), R37-R47.

Penhune, V. B., Zatorre, R. J., MacDonald, J. D., & Evans, A. C. (1996). Interhemispheric anatomical differences in human primary auditory cortex: Probabilistic mapping and volume measurement from magnetic resonance scans. *Cerebral Cortex, 6*, 661-672.

Penhune, V. B., Zatorre, R. J., & Feindel, W. H. (1999). The role of auditory cortex in retention of rhythmic patterns as studied in patients with temporal lobe removals including Heschl's gyrus. *Neuropsychologia, 37*, 315-331.

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

Peretz, I. (1996). Can we lose memory for Music? A case of music agnosia in a nonmusician. *Journal of Cognitive Neuroscience, 8*(6), 481-496.

Peretz, I., Belleville, S., & Fontaine, S. (1997). Dissociations entre musique et langage après atteinte cérébrale: un nouveau cas d'amusie sans aphasie. *Revue Canadienne de Psychologie Expérimentale, 54*(4), 354-367.

Peretz, I., Kolinsky, R., Tramo, M., Labrecque, R., Hublet, C., Demeurisse, G., & Belleville, S. (1994). Functional dissociations following bilateral lesions of auditory cortex. *Brain, 117*, 1283-1301.

Perry, D. W., & Marin, O. S. M. (1999). Neurological aspects of musical perception. In D. Deutsch (Ed.), *The Psychology of Music* (2 ed., pp. 653-724). San Diego: Academic Press.

Perry, D. W., Zattore, R. J., Petrides, M., Alivisatos, B., Meyer, E., & Evans, A. C. (1999). Localization of cerebral activity during simple singing. *NeuroReport, 10*, 3979-3984.

Petersen, S. E., Fox, P. T., Posner, M. I., Mintun, M., & al, e. (1989). Positron emission tomographic studies of the processing of single words. *Journal of Cognitive Neuroscience, 1*(2), 153-170.

Petersen, S. E., Fox, P. T., Posner, M. I., Mintun, M., & Raichle, M. E. (1988). Positron emission tomographic studies of the cortical anatomy of single-word processing. *Nature, 331*(18), 585-589.

Peterson, B., Skudlarski, P., Gatenby, J. C., Zhang, H., Anderson, A. W., & Gore, J. C. (1999). An fMRI study of Stroop word-color interference: Evidence for cingulate subregions subserving multiple distributed attentional systems. *Biological Psychiatry, 45*(10), 1237-58.

Petrides, M. (1994). Frontal lobes and working memory: evidence from investigations of the effects of cortical excisions in nonhuman primates. In F. Boller & J. Grafman (Eds.), *Handbook of Neuropsychology* (Vol. 9, pp. 59-82). New York: Elsevier Science Publishers.

Petrides, M. (1998). Working memory and the mid-dorsolateral prefrontal cortex. *Society for Neuroscience Abstracts, 24*, 18.

Petrides, M., & Pandya, D. N. (1988). Association fiber pathways to the frontal cortex from the superior temporal region in the rhesus monkey. *Journal of Comparative Neurology, 273*, 52-66.

Petrides, M., & Pandya, D. N. (1994). Comparative architectonic analysis of the human and the macaque frontal cortex. In F. Boller & J. Grafman (Eds.), *Handbook of Neuropsychology* (Vol. 9, pp. 17-58): Elsevier Science.

Petrides, M., & Pandya, D. N. (1999). Dorsolateral prefrontal cortex: comparative cytoarchitectonic analysis in the human and the macaque brain and corticocortical connection patterns. *European Journal of Neuroscience, 11*, 1011-1036.

Pitt, M. A. (1994). Perception of pitch and timbre by musically trained and untrained listeners. *Journal of Experimental Psychology: Human Perception and Performance, 20*(No. 5), 976-986.

Platt, J. R., & Racine, R. J. (1985). Effect of frequency, timbre, experience, and feedback on musical tuning skills. *Perception & Psychophysics, 38*(6), 543-553.

Raichle, M. E., Martin, W. R. W., Herscovitch, P., Mintun, M. A., & J, M. (1983). Brain blood flow measured with intravenous $H^{15}O$. II. Implementation and validation. *Journal of Nuclear Medicine, 24*, 790-198.

Rakowski, A. (1990). Intonation variants of musical intervals in isolation and in musical contexts. *Psychology of Music, 18*(1), 60-72.

Rauschecker, J. P., Tian, B., & Hauser, M. (1995). Processing of complex sounds in the macaque nonprimary auditory cortex. *Science, 268*(7), 111-114.

Risset, J.-C., & Wessel, D. L. (1999). Exploration of timbre by analysis and synthesis. In D. Deutsch (Ed.), *The Psychology of Music* (2 ed., pp. 113-169). San Diego: Academic Press.

Robin, D. A., Tranel, D., & Damasio, H. (1990). Auditory perception of temporal and spectral events in patients with focal left and right cerebral lesions. *Brain and Language, 39*, 539-555.

Romanski, L. M., Bates, J. F., & Goldman-Rakic, P. S. (1999a). Auditory belt and parabelt projections to the prefrontal cortex in the rhesus monkey. *Journal of Comparative Neurology, 403*, 141-157.

Romanski, L. M., Tian, B., Fritz, J., Mishkin, M., Goldman-Rakic, P. S., & Rauschecker, J. P. (1999b). Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. *Nature Neuroscience, 2*(12), 1131-1136.

Rossing, T. D. (1990). *The Science of Sound* (2 ed.). Reading, Massachusetts: Addison-Wesley Publishing Company.

Samson, S., & Zatorre, R. J. (1988). Melodic and harmonic discrimination following unilateral cerebral excision. *Brain and Cognition, 7*, 348-360.

Samson, S., & Zatorre, R. J. (1991). Recognition memory for text and melody of songs after unilateral temporal lobe lesion: Evidence for dual encoding. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*(4), 793-804.

Samson, S., & Zatorre, R. J. (1994). Contribution of the right temporal lobe to musical timbre discrimination. *Neuropsychologia, 32*(2), 231-240.

Samson, S., Zatorre, R. J., & Ramsay, J. O. (1997). Multidimensional scaling of synthetic musical timbre: Perception of spectral and temporal characteristics. *Canadian Journal of Experimental Psychology, 51*(4), 307-315.

Schmuckler, M. A. (1989). Expectation in music: Investigation of melodic and harmonic processes. *Music Perception, 7*(2), 109-150.

Schmuckler, M. A. (1997). Expectancy effects in memory for melodies. *Canadian Journal of Experimental Psychology, 51*(4), 292-305.

Semal, C., & Demany, L. (1991). Dissociation of pitch from timbre in auditory short-term memory. *Journal of the Acoustical Society of America, 89*(5), 2404-2410.

Semal, C., & Demany, L. (1993). Further evidence for an autonomous processing of pitch in auditory short-term memory. *Journal of the Acoustical Society of America, 94*(3), 1315-1322.

Serafine, M. L., Crowder, R. G., & Repp, B. (1984). Integration of melody and text in memory for songs. *Cognition, 16*(3), 285-303.

Serafine, M. L., Davidson, J., Crowder, R. G., & Repp, B. (1986). On the nature of melody-text integration in memory for songs. *Journal of memory and language, 25*, 123-135.

Sidtis, J. J., Sadler, A. E., & Nass, R. D. (1989). Double disconnection effects resulting from infiltrating tumors. *Neuropsychologia, 27*(11/12), 1415-1420.

Sidtis, J. J., & Volpe, B. T. (1988). Selective loss of complex-pitch or speech discrimination after unilateral lesion. *Brain and Language, 34*, 235-245.

Singh, P. G., & Hirsh, I. J. (1992). Influence of spectral locus and F0 changes on the pitch and timbre of complex tones. *Journal of the Acoustical Society of America, 92*(5), 2650-2661.

Speer, J. R., & Adams, W. E. (1985). *The effects of musical training upon the development of the perception of musical pitch.* Paper presented at the Society for Research in Child Development.

Starr, G., & Pitt, M. A. (1997). Interference effects in short-term memory for timbre. *Journal of the Acoustical Society of America, 102*(1), 486-494.

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

Stromswold, K., Caplan, D., Alpert, N., & Rauch, S. (1996). Localization of syntactic comprehension by positron emission tomography. *Brain and Language, 52*, 452-473.

Talairach, J., & Tournoux, P. (1988). *Co-planar stereotactic atlas of the human brain: 3-Dimensional proportional system: An approach to cerebral imaging.* Stuttgart and New York: Georg Thieme Verlag.

Umemoto, T. (1990). The psychological structure of music. *Music Perception, 8*(2), 115-127.

Vandenberghe, R., Price, C., Wise, R., Josephs, O., & Frackowiak, R. S. J. (1996). Functional anatomy of a common semantic system for words and pictures. *Nature, 383*(19), 254-256.

Wapnick, J., and Peter Freeman. (1980). Effects of dark-bright timbral variation on the perception of flatness and sharpness. *Journal of Research in Music Education, 28*(3), 176-184.

Warrier, C. M., Belin, P., Merlet, I., & Zatorre, R. J. (1999a). fMRI study examining effect of melodic context on pitch discrimination. *Society for Neuroscience Abstracts, 25*(Part 2), 1629.

Warrier, C. M., & Zatorre, R. J. (1999). *The role of right temporal cortex in the effect of melodic context.* Paper presented at the Society for Music Perception and Cognition, Northwestern University.

Warrier, C. M., & Zatorre, R. J. (Submitted). Influence of tonal context and timbral variation on perception of pitch. *Perception and Psychophysics.*

Warrier, C. M., Zatorre, R. J., Peretz, I., & Besson, M. (1999b). *Investigating the separation of lyrics and melody in song through PET.* Paper presented at the Brain Mapping Conference, Montreal, Quebec.

Watson, J. D. G., Myers, R., Frackowiak, R. S. J., Hajnal, J. V., Woods, R. P., Mazziotta, J. C., Shipp, S., & Zeki, S. (1993). Area V5 of the Human Brain: Evidence from a combined study using positron emission tomography and magnetic resonance imaging. *Cerebral Cortex, 3*(Mar/Apr), 79-94.

Wessel, D. L. (1979). Timbre space as a musical control structure. *Computer Music Journal, 3*, 45-52.

Whitfield, I. (1985). The role of auditory cortex in behavior. In A. Peters, and EG Jones (Ed.), *Cerebral Cortex* (Vol. 4, pp. 329-349). New York: Plenum Press.

Whitfield, I. C. (1980). Auditory cortex and the pitch of complex tones. *Journal of the Acoustical Society of America, 67*(2), 644-647.

Winstein, C. J., Grafton, S. T., & Pohl, P. S. (1997). Motor task difficulty and brain activity: Investigation of goal-directed reciprocal aiming using positron emission tomography. *Journal of Neurophysiology, 77*(3), 1581-1594.

Worsley, K. (1999). fMRIstat. Montreal: McConnell Brain Imaging Center at the Montreal Neurological Institute.

Worsley, K. J., Evans, A. C., Marrett, S., & Neelin, P. (1992). Determining the number of statistically significant areas of activation in subtracted activation studies from PET. *Journal of Cerebral Blood Flow and Metabolism, 12*, 900-918.

Worsley, K. J., Marrett, S., Neelin, P., Vandal, A. C., Friston, K. J., & Evans, A. C. (1996). A unified statistical approach for determining significant signals in images of cerebral activation. *Human Brain Mapping, 4*, 58-73.

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

Zatorre, R. J. (1988). Pitch perception of complex tones and human temporal-lobe function. *Journal of the Acoustical Society of America, 84*(2), 566-572.

Zatorre, R. J., Evans, A., Meyer, E., & Gjedde, A. (1992). Lateralization of phonetic and pitch discrimination in speech processing. *Science, 256*, 846-849.

Zatorre, R. J., Evans, A. C., & Meyer, E. (1994). Neural mechanisms underlying melodic perception and memory for pitch. *The Journal of Neuroscience, 14*(4), 1908-1919.

Zatorre, R. J., & Halpern, A. R. (1993). Effect of unilateral temporal-lobe excision on perception and imagery of songs. *Neuropsychologia, 31*(3), 221-232.

Zatorre, R. J., Halpern, A. R., Perry, D. W., Meyer, E., & Evans, A. C. (1996). Hearing in the mind's ear: A PET investigation of musical imagery and perception. *Journal of Cognitive Neuroscience, 8*(1), 29-46.

Zatorre, R. J., Mondor, T. A., & Evans, A. C. (1999). Auditory attention to space and frequency activates similar cerebral systems. *NeuroImage, 10*, 544-554.

Zatorre, R. J., Perry, D. W., Beckett, C. A., & Westbury, C. F. (1998). Functional anatomy of musical processing in listeners with absolute pitch and relative pitch. *Proceedings of the National Academy of Science, USA, 95*(March), 3172-3177.

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