Psychophysical and Neural Correlates of Auditory Attraction and Aversion

by

Kristopher Jakob Patten

A Thesis Presented in Partial Fulfillment
of the Requirements for the Degree
Master of Science

Approved October 2014 by the
Graduate Supervisory Committee:

Michael K. McBeath, Chair
Leslie C. Baxter
Eric L. Amazeen
Michael F. Dorman

ARIZONA STATE UNIVERSITY

December 2014

ABSTRACT

This study explores the psychophysical and neural processes associated with the perception of sounds as either pleasant or aversive. The underlying psychophysical theory is based on auditory scene analysis, the process through which listeners parse auditory signals into individual acoustic sources. The first experiment tests and confirms that a self-rated pleasantness continuum reliably exists for 20 various stimuli ($r = .48$). In addition, the pleasantness continuum correlated with the physical acoustic characteristics of consonance/dissonance ($r = .78$), which can facilitate auditory parsing processes. The second experiment uses an fMRI block design to test blood oxygen level dependent (BOLD) changes elicited by a subset of 5 exemplar stimuli chosen from Experiment 1 that are evenly distributed over the pleasantness continuum. Specifically, it tests and confirms that the pleasantness continuum produces systematic changes in brain activity for unpleasant acoustic stimuli beyond what occurs with pleasant auditory stimuli. Results revealed that the combination of two positively and two negatively valenced experimental sounds compared to one neutral baseline control elicited BOLD increases in the primary auditory cortex, specifically the bilateral superior temporal gyrus, and left dorsomedial prefrontal cortex; the latter being consistent with a frontal decision-making process common in identification tasks. The negatively-valenced stimuli yielded additional BOLD increases in the left insula, which typically indicates processing of visceral emotions. The positively-valenced stimuli did not yield any significant BOLD activation, consistent with consonant, harmonic stimuli being the prototypical acoustic pattern of auditory objects that is optimal for auditory scene analysis. Both the psychophysical findings of Experiment 1 and the neural processing findings of

Experiment 2 support that consonance is an important dimension of sound that is processed in a manner that aids auditory parsing and functional representation of acoustic objects and was found to be a principal feature of pleasing auditory stimuli.

ACKNOWLEDGMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Psychophysical Correlates of Auditory Attraction

The current study examines the psychophysical and neural bases for perception of sounds varying in experienced desirability. Past research has revealed numerous methods by which sounds are categorized to form hedonic judgments, ranging from the nature of the source (i.e., human versus nonhuman) (Guastavino, 2007) to higher-level emotions ascribed to the sound (e.g., rough, playful, coy) (Zentner, Grandjean, & Scherer, 2008) to characteristics of different dimensions of timbre (Herrera-Boyer, Peeters, & Dubnov, 2003; Tervaniemi, Winkler, & Naatanen, 1997). The latter method is particularly interesting as some dimensions of timbre, such as consonance, are highly quantifiable. Additionally, it is well known that preference and consonance are highly intertwined, an observation that dates back to Rameau (1722) and von Helmholtz (1877). In the current study, this relationship is first investigated psychophysically with multidimensional scaling (MDS) and then the cortical regions correlated with processing of these type of stimuli are investigated with functional magnetic resonance imaging (fMRI). Finally, the notion that this consonance/preference relationship aids in source parsing, the fundamental goal of the auditory system, is explored.

Schouten (1968) asserts that the degree of consonance in a sound is the primary dimension of timbre which, along with pitch and loudness, comprise the three fundamental aspects of sound (American Standards Association, 1960). One aim of the current study is to confirm this theory by investigating the perceptual space (via MDS) in which participants organize various environmental sounds, controlling for pitch and loudness. There have been debates on the nature of consonance as either a cultural or universal construct; the most famous

of which being that between Stumpf and Wundt over a century ago (Plomp & Levelt, 1965), though the debate remains pertinent to auditory science (Cazden, 1945; Terhardt, 1977). Recent research, however, has focused on the more quantifiable and physiologically-grounded definition of sensory consonance (McDermott & Oxenham, 2008). Sensory consonance is related to the perception of "roughness" or beat frequencies (von Helmholtz, 1877/1930) and is defined as the degree of interference of one harmonic overtone with the critical bandwidth of another (Plomp & Levelt, 1965). This interference is optimally reduced when overtones fall in line with the harmonic series (Terhardt, 1977). Due to the physics of string-and-node architecture and sound, many natural sounds tend to have spectral envelopes that approximate the harmonic series (Feynman, Leighton, & Sands, 1977; Roederer, 1995). This natural regularity may be one reason why the perception of auditory roughness and beat frequencies are considered unpleasant (Bidelman & Krishnan, 2009), and may be used by the auditory scene analysis process to parse sounds.

As defined by Bregman (1994, 2007), auditory scene analysis (ASA) is the process by which organisms use auditory regularities to parse the confluence of acoustic information reaching the ear into separate auditory objects. ASA is distinct from other theorized methods of localization and source parsing, such as using interaural time and level differences, because it is not tied to a single, occasionally fallible method. Instead, ASA exploits all natural regularities of sound sources in order to be able to localize and parse correctly even in when an acoustic signal is partially distorted (Darwin, 2008). The tendency for resonant sounds to adhere to the harmonic series, and be perceived as pleasant is one natural regularity that is likely used by ASA. Furthermore, if there are distinct neural patterns for hedonic sounds, those patterns may provide non-acoustic information that ASA can use to more efficiently parse an auditory scene.

Past research indicates that hedonic judgments are linked to degree of consonance present in sound (Bidelman & Krishnan, 2009; McDermott & Oxenham, 2008). Interestingly, preference for consonant tones is not only seen in western cultures, but also in those with different – and sometimes dissonant – musical scales (van de Geer, Levelt, & Plomp, 1962; Hatherly, 1892; Lerdahl & Jackendoff, 1983). Similarly, it is not only adults who have been inundated with musical structure and positive musical associations who exhibit this correlation; infants also tend to prefer consonant sounds (Schellenberg & Trehub, 1996a; Schellenberg & Trehub, 1996b; Trainor & Trehub, 1993a; Trainor & Trehub, 1993b). Furthermore, even hearing infants born to deaf parents who live in a deaf community prefer consonant sounds to dissonant ones (Masataka, 2006). While there is a well-documented relation between the consonant structure of sounds and human preference for them, the degree to which they are related has not been shown. The current study seeks to ascertain a measure of correlation between consonance and preference.

Neural Correlates of Auditory Attraction

Clearly, the behavioral responses to consonant and dissonant stimuli are well documented. What has been less well investigated is the neural correlates of those behavioral responses. Using functional magnetic resonance imaging (fMRI) and positron emission tomography (PET), research investigating consonant and dissonant music have found that negatively valenced stimuli elicit activation in, among other areas, the primary auditory cortex, insula, amygdala, and dorsomedial prefrontal cortex. Similarly, these studies found that positively valenced stimuli elicit activation in the primary auditory cortex, amygdala, and the anterior cingulate (Blood & Zatorre, 2001; Blood, Zatorre, Bermudez, & Evans, 1999; Koelsch, Fritz, & Schlaug, 2008; Koelsch, Fritz, Cramon, Maller, & Friederici, 2006; Koelsch, Rohrmeier, Torrecuso, & Jentschke, 2013). While the cortical activation associated with varying degrees of

consonance in musical selections have been investigated more thoroughly, the presence of musical structure may obscure the effects of consonance alone as musical genre preferences and learned associations to chord progressions are potential confounds (Johansson, 2010; Schmuckler, 1989). Pallesen and colleagues (2005) attempted to control for these confounds by using single major, minor, and dissonant chords as stimuli in an fMRI study, and found BOLD increases in similar areas when comparing minor and dissonant chords to major chords; namely, the amygdala and brain stem. The authors suggest that this pattern of activation is indicative of a cortical alarm reflex (as in Liddell, 2005), which implies that minor and dissonant chords are potentially dangerous or alarming (as proposed by Vossen, 2010).

Kumar, von Kriegstein, Friston, and Griffiths (2012) investigated neural responses to non-musical, environmental, aversive stimuli of varying degrees of dissonance. The authors propose a cognitive model of auditory processing in which stimuli are first processed in the primary auditory cortex and information travels from there to the amygdala for valence and emotional processing (see Figure 1). Using fMRI, the authors confirmed an increase in BOLD activation in the primary auditory cortex and amygdala, supporting their model. The authors also mention finding increased activation in the dorsomedial prefrontal cortex and insula, but do not include these areas in their model of auditory valence processing. From an auditory scene analysis perspective, having a neural network that assigns different valences to different auditory objects would aid in sound source segregation and localization simply by providing more information about each object. For instance, two objects near each other spatially may be segregated by valence. Additionally, following closely with Pallesen and colleagues' (2005) finding that dissonant stimuli engage a cortical alarm reflex, assignation of valences is another natural regularity that can be exploited by auditory scene analysis. The present study attempts to

expand this model by including the insula and illustrating a second network dedicated to stimulus identification, which would also aid in auditory scene analysis.
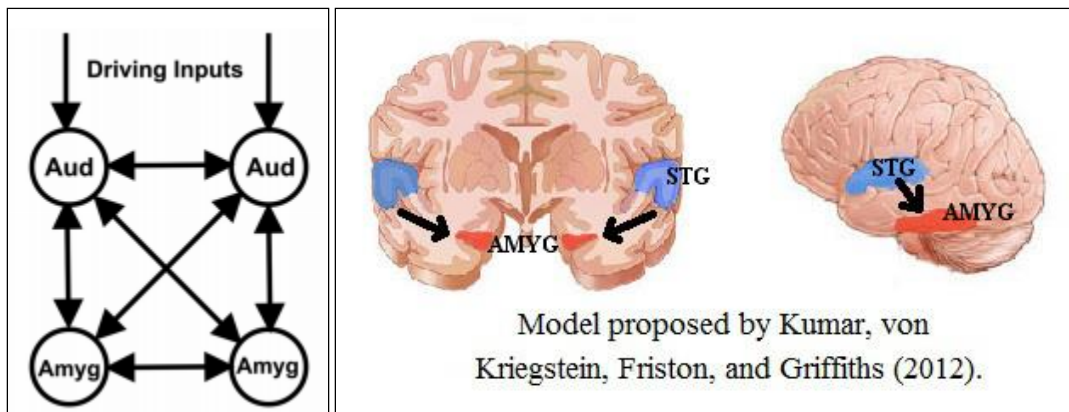

Figure 1. *Right: The model proposed by Kumar, von Kriegstein, Friston, and Griffiths (2012). This model shows that acoustic features are first processed in the primary auditory cortex, which then feeds directly to the amygdala. Left: A cythoarchitectural representation of the model. The superior temporal gyrus is shown in blue, the amygdala in red.*

Koelsch et al. (2006) noted that Heschl's gyrus, part of the primary auditory cortex located within the superior temporal gyrus, exhibited increased BOLD activation during presentation of positive stimuli. The authors attributed that increased activation to participants placing more attention on the emotionally valent auditory object, which led to greater auditory processing. This type of activation is also common in the superior temporal gyrus as a whole (Samson et al., 2011), the medial part of which comprises the rest of the primary auditory cortex (Purves, et al., 2001). Therefore, the superior temporal gyrus should exhibit greater activation in both consonant and dissonant conditions relative to the control of machine noise.

The amygdala has been implicated in processing emotional stimuli (Adolphs, Tranel, Damasio, & Damasio, 1994; LeDoux, 1992). Though not part of Kumar, von Kriegstein, Friston, and Griffiths's (2012) valence processing network model, the insula shares connectivity with the amygdala (Roy, et al., 2009; Stein, et al., 2007). Some researchers hypothesize that the insula processes the bodily reactions to emotional stimuli before sending information to the amygdala

where emotion is regulated (Craig, 2003; Paulus & Stein, 2006). The insula been identified as being active during the processing of disgusting stimuli (Kober *et al.*, 2008; Vytal *et al.*, 2010; Wicker *et al.*, 2003; Wright, He, Shapira, Goodman, & Liu, 2004). Blood and Zatorre (2001) found the insula to also be active while participants listened to extremely positive, chill-inducing music. The insula is known to be active during interoceptive processes including thirst, heartbeat regulation, and experiencing viscerally-felt emotions (Craig, 2003; Craig, 2009; Critchley, 2005; Damasio, 1999). This function of the insula accounts for activation of the structure during an intensely pleasurable musical experience and while viewing an intensely unpleasant picture. It is predicted that the insula will be more active while participants hear the strongly negatively valenced stimuli (screech). While the piano note is a positive stimuli, it is not expected to be intensely pleasurable to most participants and, thus, less likely to produce a strongly visceral response (i.e., not necessarily chills). Therefore the insula may not be strongly engaged in processing the positive stimuli.

Another region that is heavily involved in processing emotional stimuli is the dorsomedial prefrontal cortex (DMPFC). The DMPFC has been shown to increase in BOLD activation as participants view increasingly emotional pictures (Grimm, et al., 2006), but has also been implicated in conscious awareness of one's emotional state (Lane, Fink, Chua, & Dolan, 1997; Smith, et al., 2014). Furthermore, research has shown the DMPFC to be more active during tasks requiring emotional regulation and decision-making processes relating to emotional stimuli (Phillips, Drevets, Rauch, & Lane, 2003a; Phillips, Drevets, Rauch, & Lane, 2003b; Phillips, Ladouceur, & Drevets, 2008; Venkatraman, Rosati, Taren, & Huettel, 2009). In the current research, participants were asked to focus on the presented auditory stimuli and the

emotions evoked by those stimuli. As such, the DMPFC is expected to show BOLD increases in all conditions.

It is expected that participants will rate acoustic stimuli of manmade and natural origins as perceptually similar if they are of similar degrees of consonance. In other words, consonant stimuli will be rated as perceptually similar to other consonant stimuli, regardless of origin (e.g., a musical instrument or bird). It is also expected that musical instruments will be rated among the most preferential stimuli as they fall more in line with the pleasing harmonic series, and menacing animal vocalizations will be rated among the least preferential stimuli because they are dissonant in nature. Secondly, a selection of sounds ranging from very consonant to very dissonant will be presented to participants in order to determine if there are distinct patterns of activation elicited by consonant and dissonant stimuli. It is expected that all sounds will evoke increased BOLD activity in the right hemisphere as they are not speech sounds and can be differentiated mainly by pitch and timbre (Zatorre et al., 2002). Relative to baseline/control stimuli, consonant and dissonant stimuli combined will exhibit BOLD increases in the superior temporal gyrus, insula, amygdala, and dorsomedial prefrontal cortex. Relative to baseline, dissonant stimuli are expected to evoke BOLD increases in the insula, amygdala, and DMPFC. Also relative to baseline, consonant stimuli are expected to evoke increased BOLD signals in the superior temporal gyrus and DMPFC. If these areas are confirmed to be active relative to baseline, it will necessitate alteration of the model proposed by Kumar, von Kriegstein, Friston, and Griffiths (2012) to include the insula and DMPFC (see Figure 2). This new model, while still parsimonious, more fully encompasses brain regions involved in processing valenced auditory stimuli.
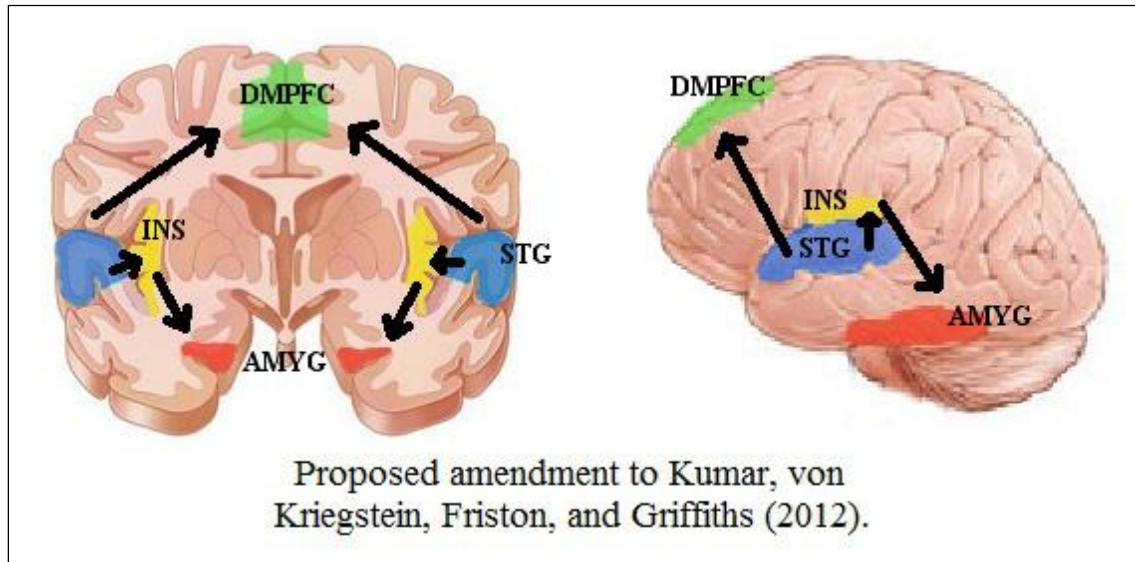
Figure 2. *The model of cortical processing of acoustic stimuli proposed by the current research. This model includes the typical auditory processing regions of the superior temporal gyrus (blue), insula (yellow), amygdala (red), and frontal areas – specifically the dorsomedial prefrontal cortex (green).*

Formal Hypotheses

Experiment 1. H1: There will be enough consistency among participants in perception of auditory similarity to obtain an interpretable multidimensional scaling result of either one or two dimensions. H2: Preference ratings for each sound will be a primary dimension by which stimuli are organized. H3: Preference ratings for each sound will be correlated with degree of consonance (measured by root mean square error from the harmonic series) present in each sound will be significantly correlated with one another.

Experiment 2. H4: Consistent with general activation of auditory areas; consonant and dissonant stimuli combined will exhibit greater blood oxygen level dependent (BOLD) activation than baseline/control stimuli in the superior temporal gyrus, insula, amygdala, and DMPFC. H5: Consistent with past findings testing negative auditory stimuli; dissonant stimuli will exhibit

greater BOLD activation than baseline in the insula, amygdala, and DMPFC. H6: Consistent

with past findings testing positive auditory stimuli: consonant stimuli will exhibit greater BOLD

activation than baseline stimuli in the superior temporal gyrus and DMPFC.

CHAPTER 2

EXPERIMENT 1

Method

Participants

Eight graduate students in a cognitive science program at Arizona State University with

no history of neurological or psychological illness participated in this experiment. Participants

had a mean age of 32.0 ($SD = 13.94$) and self-identified as non-musicians with normally

functioning hearing. Each gave informed consent in compliance with a protocol approved by the

Institutional Review Board of Arizona State University.

Stimuli and Apparatus

To accurately represent a broad variety of manmade and natural sounds, stimuli of

several different types are used; musical instruments, animal vocalizations, human vocalization,

man-made machine noise, and sounds produced without an active agent (e.g., water moving). To

control for the possible confound of differing fundamental frequencies, all stimuli were either

recorded at or digitally altered to A4 (440 Hz). For a complete list of the twenty stimuli, see

Figure 5. All stimuli are 1500 millisecond selections of a sustained sound. In this way, even the

musical instruments do not elicit a necessarily musical feel. Stimuli are audio sections captured

from high fidelity sound or video recordings. Participants listened to stimuli with Sony MDR-

7506 Studio Monitor stereo headphones. All stimuli were edited with the freeware program

Audacity and exported as wav files to avoid degradation of the sound.

Procedure

Stimuli are paired into dyads such that each sound is presented in conjunction with every

other sound. Twenty stimuli, paired, for 190 total dyads. Dyads were presented randomly to

control for possible order effects. Participants identified which of the sounds in the dyad they preferred and provided a Likert-type dissimilarity rating where 1 corresponded to "very similar" and 10 corresponded to "not similar at all".

Results

To obtain a preference metric, the number of times each sound was indicated as the preferred sound from a dyad was summed and divided by 19, the number of times each sound was presented. Using this transformation, each sound is assigned a preference rating that ranges between 0 and 1 with higher numbers corresponding to a more preferred sound. The dissimilarity ratings of the 190 dyads were entered into the Proxscal program in SPSS 22 to obtain a multidimensional scaling (MDS) solution. MDS uses a matrix of dissimilarity values to compute a map of rated stimuli in an n-dimensional space. To determine the appropriate dimensionality for a data set, Kruskal's S-stress is assessed at a series of dimensions. In common practice, the dimensionality of an MDS solution is assigned to the dimension where an "elbow" is observed in a stress (Scree) plot (Hout *et al.*, 2012; Kruskal and Wish, 1978). Instead of one obvious elbow, the data for this experiment exhibit two of what Kruskal and Wish call "weak elbows," see Figure 2. In this case, dimensionality will be attributed to the benchmarks put forth by Kruskal and Wish (1978); an appropriate dimensionality should have a stress value below .1, but anything beyond .05 may be subject to variation by random error in the data. Thus, the two dimensional solution, with a stress of .072, is appropriate for this data.
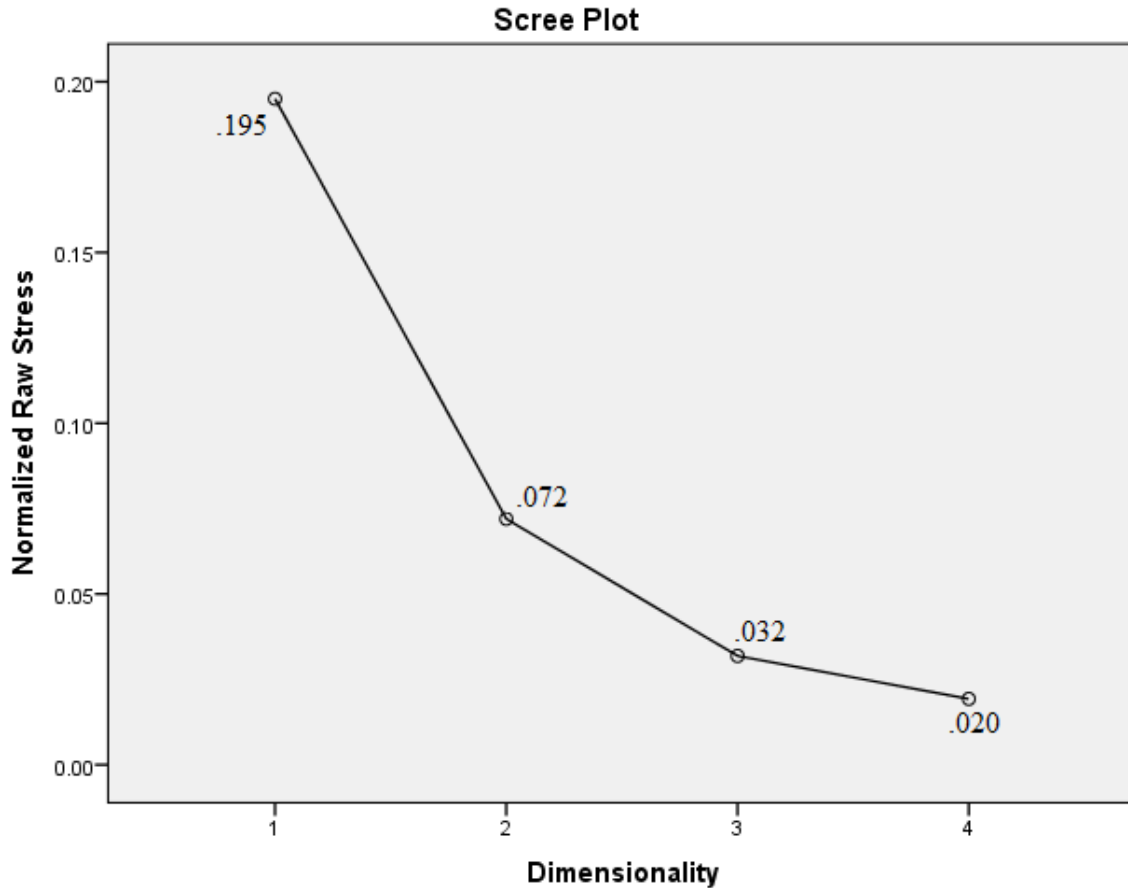
Figure 3. *A Scree plot for the sound dyad ratings. Note the two dimensional solution exhibits a stress value between .1 and .05 (Kruskal and Wish, 1978).*

MDS provides an axis-neutral plot that minimizes stress from the dissimilarity matrix. Through visual inspection, it appeared that the predominant dimension participants were using to judge dissimilarity was consonance or, as the two qualities are closely interrelated, preference. By shifting the MDS plot to align consonance along the x-axis, with the most consonant sounds in the positive direction, an x-coordinate value was assigned to each point. Additionally, the deviation of each sound's observed harmonics from its expected harmonics (given its fundamental frequency) was computed in terms of root mean square error (RMSE), as can be seen in Figure 3. Preference ratings

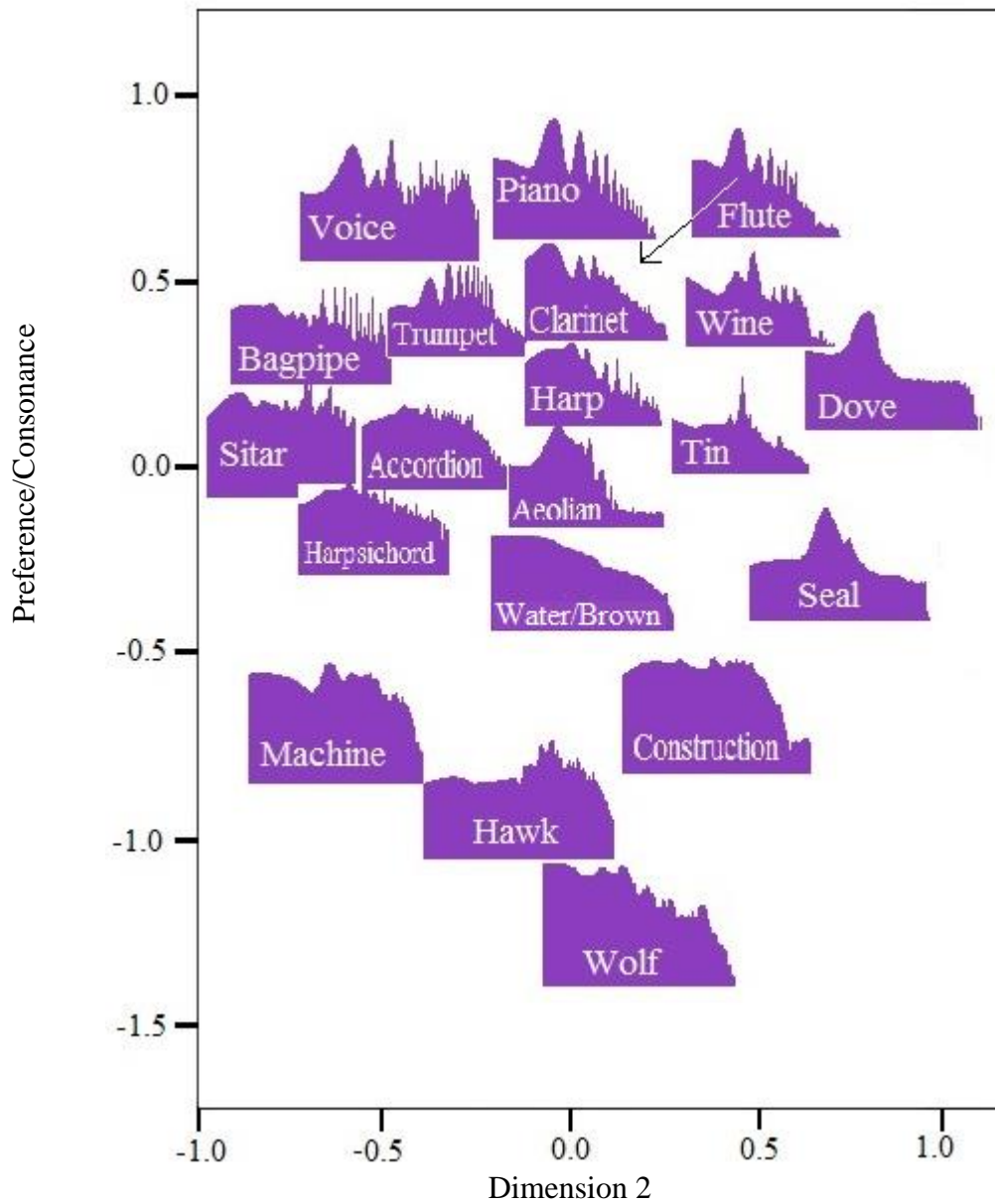Multidimensional Scaling Solution

Figure 4. *A visual representation of the MDS solution with the stimulus FFTs marking their place. Note that the flute stimulus should be located at the end of the arrow, but the FFT has been moved to show detail.*
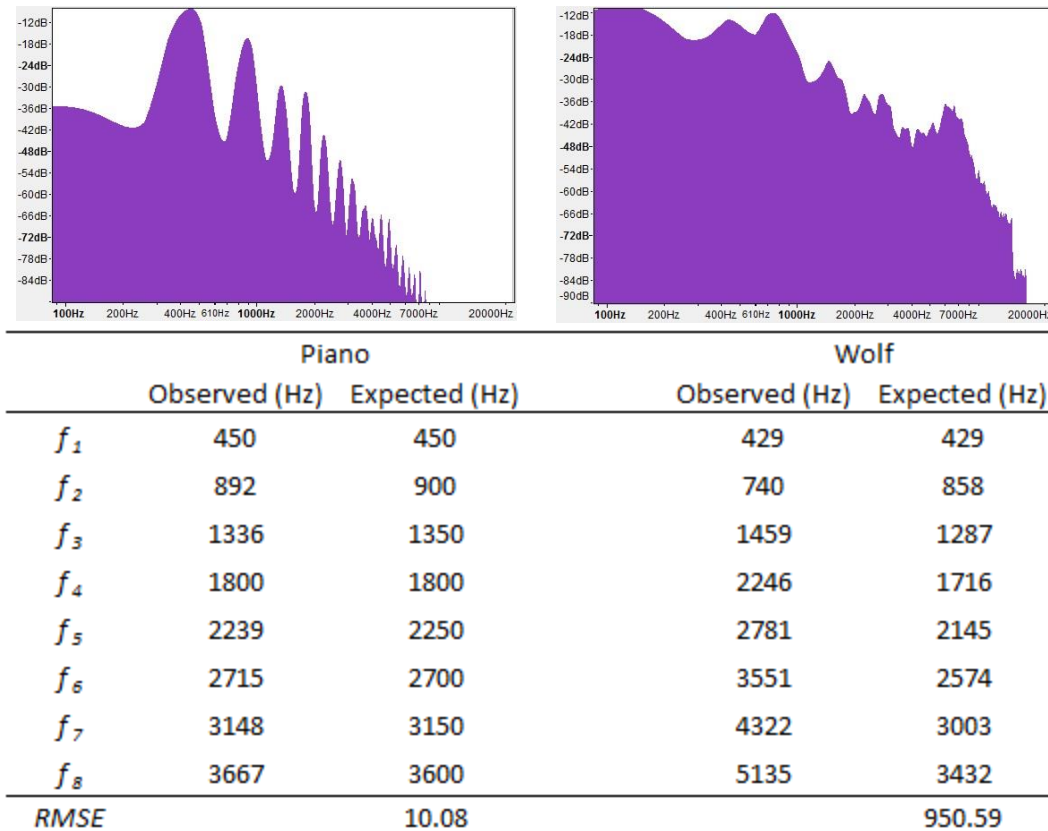
Piano			Wolf		
	Observed (Hz)	Expected (Hz)		Observed (Hz)	Expected (Hz)
f_1	450	450		429	429
f_2	892	900		740	858
f_3	1336	1350		1459	1287
f_4	1800	1800		2246	1716
f_5	2239	2250		2781	2145
f_6	2715	2700		3551	2574
f_7	3148	3150		4322	3003
f_8	3667	3600		5135	3432
RMSE		10.08			950.59

Figure 5. *Top: FFTs of single piano note and wolf growl stimuli. The piano FFT is characterized by obvious periodic peaks of energy and a low amount of inharmonic partials, while the growl is characterized by muddled, less well-defined peaks with more erratic spacing and much more inharmonic noise. Bottom: Root mean square error calculations for the two stimuli. The observed harmonics of the piano are considerably more closely aligned with those expected given its fundamental frequency. The observed fourth harmonic of the wolf growl is much closer to the expected fifth than the expected fourth; however, even when this shift is made, the wolf growl RMSE only drops to 281.73.*

are significantly negatively correlated with consonance RMSE values, $r(18) = -0.48$, $p < .05$, and positively correlated with x-coordinate values of the MDS solution, $r(18) = 0.48$, $p < .05$. Additionally, RMSE values and x-coordinates of the MDS solution were significantly negatively coordinated, $r(18) = -0.78$, $p < .001$, suggesting that consonance/preference is a valid axis for the MDS solution (Figure 4).

Discussion

The MDS solution shows that auditory events are primarily organized by preference or consonance. As previously mentioned, preference and consonance are perceptually inter-related to such a degree that the two are not easily or accurately estimated alone. The second dimension is unknown; however, the FFT data is consistent with a possible dimension of timbre, such as timbre color (the spectral envelope). The MDS is in a rough upside-down teardrop shape, with the most variance in the second dimension occurring in the more preferable sounds, suggesting that the second dimension may only be utilized for categorization of highly consonant stimuli. In this area it can be seen that sounds to the left of the midline are characterized by many high frequency harmonics with relatively large amplitudes, sounds near the middle have high frequency harmonics with lower amplitudes, and sounds to the right of the midline have one dominant frequency with few harmonics. A greater amount of harmonics typically corresponds to the sensation of a more full-bodied sound. As such, asking participants which sound they perceive as fuller or warmer in addition to which sound they prefer may shed light on the second dimension.

CHAPTER 3

EXPERIMENT 2

Method

Participants

Participants were six right-handed college graduates with a mean age of 26.67 (*SD* = 1.03). All participants self-identified as having normally functioning hearing and no history of neurological or psychological illness. Each gave informed consent in compliance with a protocol approved by the Institutional Review Boards of Arizona State University and Barrow Neurological Institute.

Stimuli

Stimuli consisted of several exemplars from Experiment 1 that were centrally located on the second dimension and spanned to preference dimension: the single piano note functioned as a positive valence exemplar and the red-tailed hawk screech and wolf growl as negative valence exemplars. In place of the rushing water sound which closely mimicked Brownian noise, actual Brownian noise was used to control for ambient impurities of the recorded sound that may have compromised the sound as a useful control. Similarly, in place of the vocal stimulus from Experiment 1, which was taken from a vocal solo by the artist Freddy Mercury, a similar vocal stimulus was performed by an amateur vocalist. This replacement was made to ensure the selection would not be recognized by any participants and skew their judgments due to strong emotional associations with the professional singer or their songs. The vocalization functioned as a second positive valence exemplar, while the Brownian noise and a pure tone sine wave functioned as control stimuli. A silent control condition was added to gauge brain function at a baseline level. As this experiment is conducted with the participants inside an fMRI machine, the

silent condition is not silence *per se*, but serves as a control for the noise of the machine (grey noise). Stimuli were presented binaurally through MRI-compatible headphones. The presentation level was adjusted to each participant's individual comfort level (between 70 and 80 dB).

Procedure

Imaging measures were conducted in the context of a classical block-design experiment, consisting of one run with a duration of 336 seconds. Each stimulus was 1500 milliseconds in duration. Stimuli were repeated eight times in a row to construct a twelve second block of constant auditory stimulation. Each twelve second block was repeated four times per stimulus in pseudorandom order (i. e., no two same blocks in a row) for a total of twenty-eight individual blocks. Participants were instructed to close their eyes to limit visual cortex activation and to allow better focus on the auditory stimuli. Additionally, each participant was asked to listen to the sounds and focus on how it made them feel, rather than sound quality, source, or other aspects of the stimuli. Stimuli were delivered into the scanner via a Sony Vaio laptop computer connected to an MR-compatible headphone system from Nordic Neurolab (Bergan, Norway). The software Presentation (Neurobehavioral Systems, Inc.) was used for delivering the auditory stimuli. Participants were also fitted with a foam head surround to reduce movement in the scanner. After exiting the fMRI machine, participants were instructed to rate the sounds they heard on a 100 point Likert-type scale, with 100 corresponding to a perfect, intensely pleasurable sound and 0 corresponding to a terrible sound.

Data Acquisition and Analysis

Functional magnetic resonance imaging (fMRI) was performed on a GE 3-Tesla MR scanner. T_2* weighted images were acquired with a gradient echo, echo-planar pulse sequence to elicit (BOLD) contrast. The scanning parameters were as follows: TE (echo time) = 20 ms; TR

(repetition time) = 3.0 s; flip angle = 80°; acquisition matrix = 64 x 64 voxels; field of view

(FOV) = 240 mm, 4 mm slice thickness with no gap. 40 slices of the brain were acquired axially

within the TR at each time point, with near isotropic voxel resolution of 3.75 x 3.75 x 4.0 mm.

High resolution structural images were acquired for normalization of the results. The

images were collected using a 3D SPGR (spoiled gradient) T_1-weighted, 3D acquisition with the

following parameters: TR = 24 ms; TE = 6 ms; flip angle = 40°; NEX = 1; slice thickness = 1.0

mm; 0 skip between slices; FOV = 24 cm; in-plane resolution = 0.9375 mm^2 voxels. The T_1 and

T_2* weighted images were co-registered using a least squares minimization routine.

The statistical parametric mapping software package (SPM8, Wellcome Department of

Cognitive Neurology, London, UK) implemented in Matlab was used for the preprocessing and

statistical analyses. The time-series of images were corrected for motion, normalized into a

standard atlas space (using the International Consortium for Brain Mapping template as

implemented in SPM8), and then spatially smoothed using an 8 mm full-width at half-maximum

Gaussian kernel. Individual time-series analysis was performed on each participant.

Regions of interest (ROI) masks defined by the specific hypothesis were obtained by selecting

relevant brain regions from the WFU Pickatlas toolbox for SPM (Version 2.4, Lancaster,

Summerln, Rainey, Freitas, & Fox, 1997; Lancaster et al., 2000; Maldjian, Laurienti, Burdette, &

Kraft, 2003; Maldjian Laurienti, & Burdette, 2004). Each resulting ROI was corrected for

family-wise error (FWE). Hemodynamic data from the ROI analysis were further visualized in

the form of peristimulus time histograms. These histograms were created with the rfxplot toolbox

for SPM (Glascher, 2009).

Results

Baseline Stimuli

As previously mentioned, the "silent" control condition is not silent as the noise of the fMRI machine is constant throughout the experiment. Additionally, post-imaging stimulus ranking revealed that participants perceived the pure tone sine wave as more positive than neutral (rescaled $M = 76.11$, $SD = 22.00$), suggesting that it was not a useful control. Brownian noise, on the other hand was rated much more moderately (rescaled $M = 41.39$, $SD = 4.71$). Similarly, because the machine noise was present during all conditions, it was not a true baseline measure and was not used as such. To ensure the data obtained from the experimental contrasts was not spurious, we also analyzed a region not expected to show differences in activation based on the stimuli. As a control region, we chose to examine the calcarine cortex, which is involved in visual functioning and shows strong activation to visual stimuli. As expected, this region did not yield suprathreshold activation on any of the comparisons.

Stimulus Ratings

After exiting the scanner, all participants were asked to indicate the degree to which they preferred each stimulus on a scale from 0 to 100, with higher numbers corresponding to a greater preference. Interestingly, ratings from all participants were weighted more towards the negative spectrum; ratings ranged from 2.2 ($SD = 2.71$) to 72 ($SD = 8.12$). As the original scale was comprised of arbitrary units, proportionally rescaling the ratings to fit a 100 point scale within the 72 point range does not change the relationship between stimuli, but does make those relationships easier to visualize. For instance, the Brownian noise stimuli, which held a position between consonant and dissonant stimuli in the two dimensional scaling solution in Experiment 1 received a rating of 29.8 ($SD = 4.71$) on the original scale. While this is not far from the

halfway point between the lower range of the scale and the highest rated stimuli (36), rescaling

this value to 41.39 is easier to visualize as a point near a round number midpoint (50). Figure 4

shows the average participants ratings and standard error of each stimulus. It is also of note that,

aside from some sine and vocal stimuli, the rank ordering of all six stimuli were the same among
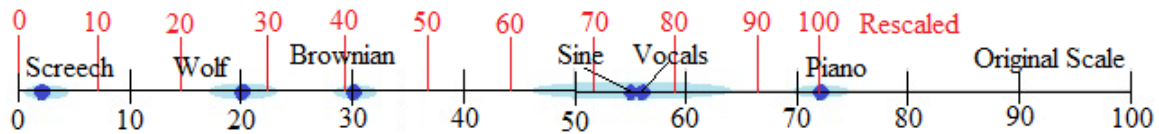
all participants.

Figure 6. *Ratings of experimental and control stimuli on both original and rescaled planes. The light blue ovals surrounding each point represent the standard error.*

Experimental Stimuli

Comparing all experimental stimuli to the baseline stimulus (All>Brown) yielded

significant BOLD increases in the right superior temporal gyrus (SPM $T = 14.43$, 461 voxels,

corrected $p = .013$), left superior temporal gyrus (SPM $T = 9.43$, 278 voxels, corrected $p = .036$),

and left dorsomedial prefrontal cortex (SPM $T = 13.21$, 36 voxels, corrected $p = .003$). Images of

these regions are presented in Figure 5. The coordinates of these regions can be found in Table 1.
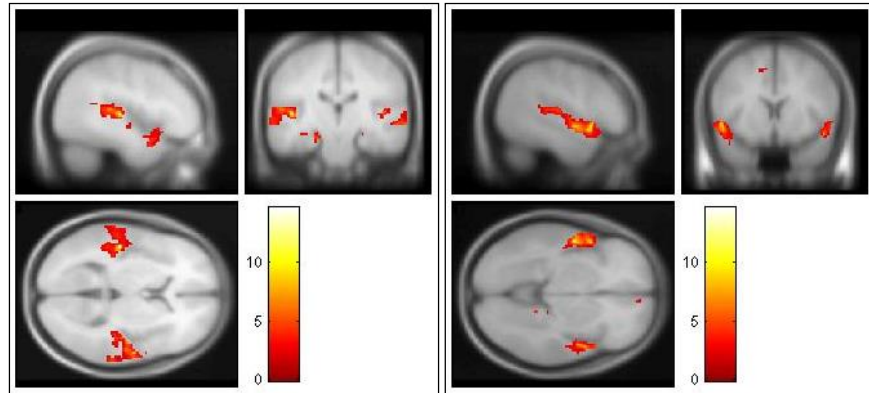
STG
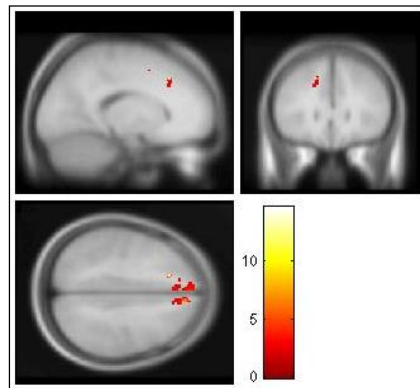
DMPFC

Figure 7. *BOLD response from the three regions significantly more active across all experimental stimuli compared to the Brownian noise baseline. Top left and right: bilateral superior temporal gyrus. Bottom: left dorsomedial prefrontal cortex (right dorsomedial prefrontal cortex is also pictured, but does not reach suprathreshold levels).*
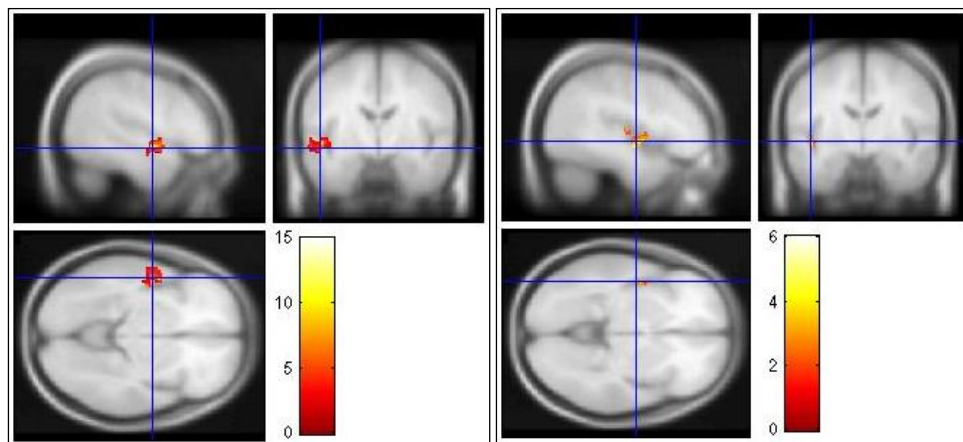
STG   *Insula*

Figure 8. *BOLD response from the two regions significantly more active when participants listened to the screech stimuli in comparison to the Brownian noise baseline. Left: left anterior portion of the superior temporal gyrus. Right: left insula.*

STG   Insula
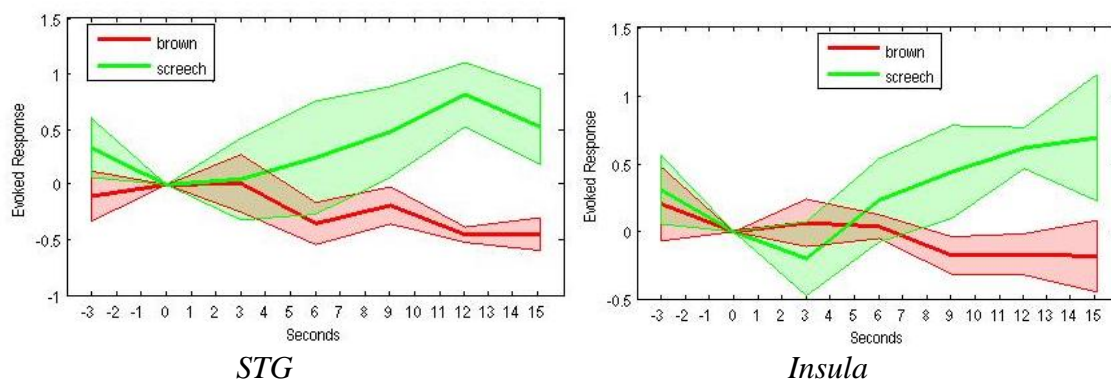
Figure 9. *Peristimulus time histograms of the two regions significantly more active when participants listened to the screech stimuli in comparison to the Brownian noise baseline. X-axes for these figures are unitless values that show percent change of the blood oxygen level dependent signal. Shaded regions surrounding the solid lines are standard errors. Figures have been scaled to zero at the stimulus onset. Left: left superior temporal gyrus. Right: left insula.*

The participants varied somewhat on how they emotionally interpreted the experimental stimuli, although there was good agreement in rank, indicating that the screech and piano were universally the least and most preferred. We performed separate analyses on these two stimuli that also showed the greatest separation in preference; the negatively valenced hawk screech (rescaled $M = 3.06$, $SD = 2.71$) and the positively valenced piano note (rescaled $M = 100$, standard deviation = 8.12). Despite the reliably high rating of the piano stimulus (Piano>Brown), it did not elicit any significant BOLD activation above baseline. The red-tailed hawk screech (Screech>Brown), however, elicited significant activation above baseline in the left superior temporal gyrus (SPM $T = 14.91$, 296 voxels, FWE corrected $p = .03$) and activation of marginal significance in the left insula (SPM $T = 4.48$, 78 voxels, FWE corrected $p = .075$). Images of these regions are presented in Figures 6 and 7, coordinates are presented in Table 1.

Table 1
Coordinates of Clusters of Activation Showing Significantly Different BOLD Activity for
Auditory Stimuli in Comparison to Brownian Noise Control

Region	k	x	y	z	BA
All>Brown					
Right superior temporal gyrus	461	58	-18	-2	21
Left superior temporal gyrus	278	-50	10	-6	22
Left dorsomedial prefrontal cortex	36	-16	28	42	8
Screech>Brown					
Left superior temporal gyrus	296	-50	-6	-8	21
Left insula	78	-44	-10	-8	21
Piano>Brown					
No significant clusters of activation					

Note - All clusters reached an uncorrected significance level of $p = .01$ and an extent threshhold of 20 voxels. For each cluster, coordinates are given for the maximally activated voxel. k, number of voxels in cluster; BA, Brodmann's areas; x, y, z, MNI coordinates.

Discussion

Comparing all stimuli to the Brownian noise (All>Brown) revealed significant activation in several areas. Among these were the bilateral superior temporal gyri; activation of these areas suggests that the experimental stimuli captured auditory attention (Ahveninen *et al*., 2011; Jancke, Mirzazade, & Joni Shah, 1999) and, as such, may have been more deeply processed. Additionally, significant activation was also observed in the left DMPFC, an area that has been identified as playing a role in conscious perception of emotion, emotion regulation, and decisions regarding emotional content (Grimm, et al., 2006; Lane, Fink, Chua, & Dolan, 1997; Phillips, Ladouceur, & Drevets, 2008; Venkatraman, Rosati, Taren, & Huettel, 2009). Activation in this area suggests that the participants attended to the auditory stimuli and found them at least partly emotionally evocative. It is worth noting that, while activation was detected in this area when comparing all experimental stimuli to the baseline, it was not detected upon comparing the extremes to the baseline. This may be due to the decreased power of detection when examining

the smaller single blocks rather than collapsing the experimental stimuli together. In other words, this DMPFC activation may be present in each condition, but cannot be detected for single stimuli alone with this data set.

Relative to baseline, the hawk screech elicited increased BOLD activation in the left superior temporal gyrus and the left insula, a finding which supports the proposed model of cortical regions involved in processing valenced acoustic stimuli. Both the baseline brown noise stimulus and the negative-valence hawk screech stimulus were presented at identical intensity levels; thus, the increased activation detected in auditory regions is not due to a stronger acoustic signal in the screech stimuli. Increased activation in the superior temporal gyrus may signify increased processing of the hawk screech in the primary and secondary cortex associated with auditory processing. Previous studies have shown similar activation of auditory cortex in response to aversive sounds (Kumar, von Kriegstein, Friston, & Griffiths, 2012) and arousing pleasant sounds (Blood & Zatorre, 2001; Koelsch et al., 2006; Koelsch et al., 2013). This suggests that, in the current study, the screech sound was more arousing (in a negative way) than the piano, which aligns with the behavioral ratings.

The negative valence stimulus also produced activation in the insula. Blood and Zatorre (2001) and Koelsch et al. (2006) found the insula to be more active during positive stimulus presentation in relation to a neutral baseline and Kumar, von Kriegstein, Friston, and Griffiths (2012) found the insula to be more active during presentation of highly dissonant, negatively valenced stimuli. Additionally, other studies have found the insula to be active in the processing of fear and disgust (Calder, 2003; Schienle et al., 2005; Zeki & Romaya, 2008). These are emotions that are experienced viscerally and interoceptively (Adolphs, Tranel, & Damasio, 2003; Craig, 2009; Damasio, 1999; Damasio, 2003). The finding of the current study that the hawk

screech elicits strong activity in the insula relative to baseline is in concert with these findings; the highly dissonant screech likely elicits a strong, visceral emotion of fear or disgust. Indeed, some participants related colloquially that they found it difficult to avoid flinching when the screech was presented.

Both the main effect of all sounds versus baseline, as well as the specific findings for the negative stimuli vs. baseline, showed a high degree of left-lateralized activity. In contrast, many studies that have found right-lateralized responses to stimuli (Blood & Zatorre, 2001; Blood, Zatorre, Bermudez, & Evans, 1999; Pallesen, et al., 2005). While past research has held that the right hemisphere is dominant in tone processing and the left hemisphere in speech processing (Belin, et al., 1998; Zatorre, Belin, & Penhune, 2002), more recent studies call that into question. Scheich, Brechmann, Bosch, Budinger, and Ohl (2007) noted that the auditory task given to participants was more indicative of hemispheric differences than the stimuli presented to them. This finding has been interpreted as an example of local (as opposed to global) auditory processing, which encompasses comparative, relational, and analytic tasks (Bever & Chiarello, 1974; Justus & A, 2005; Rauschecker, Tian, Pons, & Mishkin, 1997; Wetzel, Ohl, & Scheich, 2008). Moreover, categorization of auditory stimuli is also a function of local processing (Brechmann & Scheich, 2005; Brechmann, et al., 2007). The dorsomedial prefrontal cortex was observed to be more active during perception of valenced sounds than a neutral baseline, which is indicative of stimulus identification and categorization. As left lateralization held true for both all valenced stimuli and negatively valenced stimuli, it is likely that the dorsomedial prefrontal cortex may also be more active during perception of negatively valenced sounds compared to baseline, but was not able to reach significance.

The stimulus rated as most positive did not elicit any significant change in BOLD activity relative to baseline (Piano>Brown) while the negative stimulus did. Behaviorally, the negative stimulus elicited a stronger negative response than did the piano elicit a positive response. This difference may be due to arousal levels evoked by each stimulus. While the single piano note is a reliably positive stimulus, it likely does not evoke as strong a positive emotion as music. Blood and Zatorre's (2001) found that music that is characterized by higher levels of arousal in the positive valence (defined in their study as evoking a sensation of chills) produces increased BOLD activity in several brain regions compared to music that does not arouse participants in the same way. It is probable that a positively valenced stimulus that is evokes low levels of arousal is experienced more as a calming event than an intensely pleasurable one. The International Affective Digitized Sounds database contains short musical selections of several musical instruments (Bradley & Lang, 2007). While these sounds were rated as mostly pleasurable, they were not rated as arousing. In fact, Bradley and Lang's 100 participants rated a harp among the most pleasurable sounds (rating = 7.44, range = 1.57 – 7.9), but as one of the least arousing sounds (rating = 3.36, range = 3.36 – 8.16).

CHAPTER 4

GENERAL DISCUSSION

The first experiment demonstrates that sensory consonance, a universal perception, correlates with participants' preference for that sound. This suggests that preference is primarily stable, universal, and directly correlated to consonance. Moreover, it establishes consonance, along with pitch and loudness, as a primary aspect of sound used in categorization and parsing. As one of the primary dimensions of sound categorization, degree of consonance is likely a natural regularity used in auditory scene analysis. The first experiment also illustrates the close proximity of consonant natural sounds and consonant manmade sounds, such as musical instruments. While it may be a natural assumption to group musical instruments in a separate category from other sounds (Guastavino, 2007), the results of Experiment 1 suggest that this is a false dichotomy; all sounds lie along the same two dimensional plane. Furthermore, the two dimensional scaling solution also shows that musical instruments appear to extend beyond consonant natural sound in the preference dimension. This could be evidence that early humans sought to refine, replicate, and domesticate the non-arousing and prototypical consonant stimuli that were hallmarks of calm, safe locations.

The second experiment illustrates the neural regions that are involved in processing of emotion-evoking sounds. In accordance with other research investigating cortical processing of highly arousing sounds, the auditory cortices and insula were identified as exhibiting greater BOLD activation relative to a baseline condition. Notably, only the negatively valenced (and potentially highly arousing) sound elicited activation in these areas, while the positively valenced (and potentially soothing) sound elicited no activation compared to a neutral baseline sound. These findings also fall in line with an evolutionary hypothesis of music development; highly

arousing sounds of both positive and negative valences would prompt humans to act, while non-arousing, positively valenced stimuli would indicate a safe area and promote calm emotions. Highly negative and highly arousing stimuli, like the hawk screech, indicate danger and prime a sympathoadrenal, fight or flight response. Highly positive and highly arousing stimuli, such as erotic moans or laughter (Bradley & Lang, 2007), indicate social interaction and prime an appropriate social response. A stimulus that is highly positive and not arousing, like the piano note, indicate safety but do not necessitate any response. It may be that humans have come to expect non-arousing positive stimuli in their daily life (e. g., bird songs, music in parking lots, calm conversations in restaurants). Indeed, the physics of resonant systems, such as string and node or pipe and antinode, dictate that most sounds have spectral energy that closely approximates the harmonic series (Feynman, Leighton, & Sands, 1977; Roederer, 1995). Because of this, non-arousing positively valenced stimuli are treated as a type of baseline against which to measure arousing and/or negative stimuli. This notion is also supported by the multidimensional solution from Experiment 1. Positively rated consonant sounds showed a greater amount of variability along the second (possibly tone color) axis than the negatively rated dissonant sounds, which may indicate the second dimension becoming less important for categorization as sound becomes more dissonant.

The results showed that consonance is highly correlated with preference and that preference/consonance is a primary dimension by which sound is categorized. Additionally, the auditory cortices, insula, and DMPFC were shown to be more active during perception of valenced sounds than during perception of a baseline stimulus. This suggests that the auditory cortex/amygdala model of emotional auditory perception proposed by Kumar, von Kriegstein, Friston, and Griffiths (2012) should be amended to include the insula and DMPFC. The low

power obtained when comparing soothing positive sounds to baseline suggest that these sounds may be more ecologically valid as baseline stimuli than common stimuli like white noise or silence. As such, future psychophysical and imaging research may benefit by using a negligibly-arousing, consonant stimulus as a secondary control, if not a primary. Using the correct control in such studies can reveal more correct and more detailed information than a flawed control. Finally, the reliable activation of cortical areas in response to dissonant sounds is a further natural regularity – beyond degree of consonance – than can aid in auditory scene analysis. In fact, the pattern of cortical activation in response to dissonant sounds is particularly adaptive, as it replaces the regularity of consonance that is removed when a sound deviates from the harmonic series.

REFERENCES

Adolphs, R., Tranel, D., & Damasio, A. R. (2003). Dissociable neural systems for recognizing emotions. *Brain and cognition, 52*(1), 61-69.

Adolphs, R., Tranel, D., Damasio, H., & Damasio, A. (1994). Impaired recognition of emotion in facial expressions following bilateral damage to the human amygdala. *Nature, 372*, 669-672.

Ahveninen, J., Hamalainen, M., Jaaskelainen, I. P., Ahlfors, S. P., Huang, S., Lin, F. H., . . . Belliveau, J. W. (2011). Attention-driven auditory cortex short-term plasticity helps segregate relevant sounds from noise. *Proceedings of the National Academy of Sciences, 108*(10), 4182-4187.

American Standards Association. (1960). *American standard acoustical terminology.* New York: American Standards Association.

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

Bidelman, G., & Krishnan, A. (2009). Neural correlates of consonance, dissonance, and the hierarchy of musical pitch in the human brainstem. *The Journal of Neuroscience, 29*(42), 13165-13171.

Blood, A. J., & Zatorre, R. J. (2001). Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proceedings of the National Academy of Sciences of the United States of America, 98*(20), 11818-11823.

Blood, A. J., Zatorre, R. J., Bermudez, P., & Evans, A. C. (1999). Emotional responses to pleasant and unpleasant music correlate with activity in paralimbic brain regions. *Nature Neuroscience, 2*(4), 382-387.

Bradley, M. M., & Lang, P. J. (2007). The International Affective Digitized Sounds (IADS-2): Affective ratings of sounds and instruction manual. *University of Florida, Gainesville, FL, Technical Report B-3*.

Brechmann, A., & Scheich, H. (2005). Hemispheric shifts of sound representation in auditory cortex with conceptual listening. *Cerebral Cortex, 15*, 578-587.

Brechmann, A., Gaschler-Markefski, B., Sohr, M., Yoneda, K., Kaulisch, T., & Scheich, H. (2007). Working memory–specific activity in auditory cortex: potential correlates of sequential processing and maintenance. *Cerebral Cortex, 17*(11), 2544-2552.

Bregman, A. S. *Auditory Scene Analysis: The perceptual organization of sound*. MIT press, 1994.

Bregman, A. S. (2007, October). Progress in the study of auditory scene analysis. In *Applications of Signal Processing to Audio and Acoustics, 2007 IEEE Workshop* (pp. 122-126). IEEE.

Calder, A. J., Keane, J., Manly, T., Sprengelmeyer, R., Scott, S., Nimmo-Smith, I., & Young, A. W. (2003). Facial expression recognition across the adult life span. *Neuropsychologia, 41*(2), 195-202.

Cazden, N. (1945). Musical consonance and dissonance: A cultural criterion. *The Journal of Aesthetics and Art Criticism, 4*(1), 3-11.

Craig, A. (2003). Interoception: the sense of the physiological condition of the body. *Current Opinions in Neurobiology, 13*(4), 500-505.

Craig, A. (2009). How do you feel - now? The anterior insula and human awareness. *Nature Reviews Neuroscience, 10*(1), 59-70.

Critchley, H. D. (2005). Neural mechanisms of autonomic, affective, and cognitive integration. *Journal of Comparative Neurology, 493*(1), 154-166.

Damasio, A. (2003). Feelings of emotion and the self. *Annals of the New York Academy of Sciences, 1001*(1), 253-261.

Damasio, A. R. (1999). *The Feeling of What Happens: Body and Emotion in the Making of Consciousness.* New York: Harcourt Brace.

Darwin, C. J. (2008). Spatial hearing and perceiving sources. In W. A. Yost, A. N. Popper, & R. R. Fay, *Auditory Perception of Sound Sources* (pp. 215-232). New York: Springer.

Feynman, R. P., Leighton, R. B., & Sands, M. (1977). *The Feynman Lectures of Physics, Volume 1.* Boston: Addison-Wesley.

Glascher, J. (2009). Visualization of group inference data in functional neuroimaging. *Neuroinformatics, 7*(1), 73-82.

Grimm, S., Schmidt, C. F., Bermpohl, F., Heinzel, A., Dahlem, Y., Wyss, M., . . . Northoff, G. (2006). Segregated neural representation of distinct emotion dimensions in the prefrontal cortex - an fMRI study. *NeuroImage, 30*(1), 325-340.

Guastavino, C. (2007). Categorization of environmental sounds. *Canadian Journal of Experimental Psychology, 61*(1), 54-63.

Hatherly, S. G. (1892). *A Treatise on Byzantine Music.* London: A. Gardner.

Herrera-Boyer, P., Peeters, G., & Dubnov, S. (2003). Automatic classification of musical instrument sounds. *Journal of New Music Research, 32*(1), 3-21.

Hout, M. C., Papesh, M. H., & Goldinger, S. D. (2012). Multidimensional scaling. *Wiley Interdisciplinary Reviews (WIREs): Cognitive Science, 4*, 93-103.

Jancke, L., Mirzazade, S., & Joni Shah, N. (1999). Attention modulates activity in the primary and the secondary auditory cortex: A functional magnetic resonance imaging study in human subjects. *Neuroscience Letters, 266*(2), 125-128.

Johansson, K. G. (2010). What chord was that? A study of strategies among ear players in rock music. In C. F. Thorgersen, & S. Karlsen, *Music, education and innovation* (pp. 17-32). Pitea, Sweden: University Press, Lulea University of Technology.

Justus, T., & A, L. (2005). Auditory attention to frequency and time: An analogy to visual local-global stimuli. *Cognition, 98*, 31-51.

Kober, H., Barrett, L. F., Joseph, J., Bliss-Moreau, E., Lindquist, K., & Wager, T. D. (2008). Functional grouping and cortical-subcortical interactions in emotion: A meta-nalysis of neuroimaging studies. *Neuroimage, 42*(2), 998-1031.

Koelsch, S., Fritz, T., & Schlaug, G. (2008). Amygdala activity can be modulated by unexpected chord functions during music listening. *Neuroreport, 19*(18), 1815-1819.

Koelsch, S., Fritz, T., Cramon, D., Maller, K., & Friederici, A. (2006). Investigating emotion with music: An fMRI study. *Human Brain Mapping, 27*(3), 239-250.

Koelsch, S., Rohrmeier, M., Torrecuso, R., & Jentschke, S. (2013). Processing of hierarchical syntactic structure in music. *Proceedings of the National Academy of Sciences, 110*(38), 15443-15448.

Kruskal, J. B., & Wish, M. (1978). *Multidimensional Scaling. Sage University Paper Series on Quantitative Application in the Social Sciences.* Beverly Hills and London: Sage Publications.

Kumar, S., von Kriegstein, K., Friston, K., & Griffiths, T. D. (2012). Features versus feelings: Dissociable representations of the acoustic features and valence of aversive sounds. *The Journal of Neuroscience, 32*(41), 14184-14192.

Lancaster, J. L., Rainey, L. H., Summerlin, J. L., Freitas, C. S., Fox, P. T., Evans, A. C., . . . Mazziotta, J. C. (1997). Automated labeling of the human brain: A preliminary report on the development and evaluation of a forward-transform method. *Human Brain Mapping, 5*(4), 238.

Lancaster, J. L., Woldorff, M. G., Parsons, L. M., Liotti, M., Freitas, C. S., Rainey, L., . . . Fox, P. T. (2000). Automated Talairach atlas labels for functional brain mapping. *Human Brain Mapping, 10*(3), 120-131.

Lane, R., Fink, G., Chua, P., & Dolan, R. (1997). Neural activation during selective attention to subjective emotional responses. *Neuroreport, 8*(18), 3969-3972.

LeDoux, J. E. (1992). Emotion and the amygdala. In J. P. Aggleton, *The amygdala: Neurobiological aspects of emotion, memory, and mental dysfunction* (pp. 339-351). New York: Wiley.

Lerdahl, F., & Jackendoff, R. (1983). *A Generative Theory of Tonal Music.* Cambridge, Massechusetts: MIT Press.

Liddell, B. J., Brown, K. J., Kemp, A. H., Barton, M. J., Das, P., Peduto, A., . . . Williams, L. M. (2005). A direct brainstem-amygdala-cortical-alarm-system for subliminal signals of fear. *NeuroImage, 24*(1), 235-243.

Maldjian, J. A., Laurienti, P. J., & Burdette, J. H. (2004). Precentral gyrus discrepancy in electronic versions of the Talairach atlas. *Neuroimage, 21*(1), 450-455.

Maldjian, J. A., Laurienti, P. J., Burdette, J. B., & Kraft, R. A. (2003). An automated method for neuroanatomic and cytoarchitectonic Atlas-based interrogation of fMRI data sets. *Neuroimage, 19*, 1233-1239.

Masataka, N. (2006). Preference for consonance over dissonance by hearing newborns of deaf parents and of hearing parents. *Developmental Science, 9*(1), 46-50.

McDermott, J. H., & Oxenham, A. J. (2008). Music perception, pitch, and the auditory system. *Current Opinions in Neurobiology, 18*(4), 452-463.

Pallesen, K. J., Brattico, E., Bailey, C., Korvenoja, A., Koivisto, J., Gjedde, A., & Carlson, S. (2005). Emotion processing of major, minor, and dissonant chords. *Annals of the New York Acadmey of Sciences, 1060*, 450-453.

Paulus, M. P., & Stein, M. B. (2006). An insular view of anxiety. *Biological Psychiatry, 60*(4), 383-387.

Peretz, I., Blood, A. J., Penhune, V., & Zatorre, R. (2001). Cortical deafness to dissonance. *Brain, 124*, 928-940.

Peretz, I., Brattico, E., Jarvenpaa, M., & Tervaniemi, M. (2009). The amusic brain: In tune, out of key, and unaware. *Brain, 132*(5), 1277-1286.

Phillips, M. L., Drevets, W. C., Rauch, S. L., & Lane, R. (2003a). Neurobiology of emotion perception I: The neural basis of normal emotion perception. *Biological Psychiatry, 54*(5), 504-514.

Phillips, M. L., Drevets, W. C., Rauch, S. L., & Lane, R. (2003b). Neurobiology of emotion perception II: Implications for major psychiatric disorders. *Biological Psychiatry, 54*(5), 515-528.

Phillips, M. L., Ladouceur, C. D., & Drevets, W. C. (2008). A neural model of voluntary and automatic emotion regulation: Implications for understanding the pathophysiology and neurodevelopment of bipolar disorder. *Molecular Psychiatry, 13*(9), 833-857.

Plomp, R., & Levelt, W. J. (1965). Tonal consonance and critical bandwidth. *The Journal of the Acoustical Society of America, 38*, 548-560.

Purves, D., Augustine, G. J., Fitzpatrick, D., Hall, W. C., LaMantia, A. S., & White, L. E. (2001). *Neuroscience.* Sunderland, MA: Sinauer Associates.

Rameau, J. P. (1722/1971). *Treatise on Harmonics.* New York: Dover Publications.

Rauschecker, J. P., Tian, B., Pons, T., & Mishkin, M. (1997). Serial and parallel processing in rhesus monkey auditory cortex. *Journal of Computational Neurology, 382*, 89-103.

Roederer, J. G. (1995). *The Physics and Psychophysics of Music: An Introduction.* New York: Springer.

Roy, A. K., Shehzad, Z., Margulies, D. S., Kelly, A. M., Uddin, L. Q., Gotimer, K., . . . Milham, M. P. (2009). Functional connectivity of the human amygdala using resting state fMRI. *Neuroimage, 45*(2), 614-626.

Samson, F., Zeffiro, T. A., Toussaint, A., & Belin, P. (2011). Stimulus complexity and categorical effects in human auditory cortex: An activation likelihood estimation meta-analysis. *Frontiers in Psychology, 1*, 241.

Scheich, H., Brechmann, A., Brosch, M., Budinger, E., & Ohl, F. W. (2007). The cognitive auditory cortex: Task-specificity of stimulus representations. *Journal of Hearing Research, 229*, 213-224.

Schellenberg, E. G., & Trehub, S. (1996b). Children's discrimination of melodic intervals. *Developmental Psychology, 32*, 1039-1050.

Schellenberg, E. G., & Trehub, S. E. (1996a). Natural musical intervals: Evidence from infant listeners. *Psychological Science, 7*(5), 272-277.

Schienle, A., Stark, R., Walter, B., Blecker, C., Ott, U., Kirsch, P., . . . Vaitl, D. (2002). The insula is not specifically involved in disgust processing: An fMRI study. *Neuroreport, 13*(16), 2023-2026.

Schmuckler, M. A. (1989). Expectations in Music: Investigation of Melodic and Harmonic Processes. *Music Perception, 7*(2), 109-150.

Schouten, J. F. The perception of timbre. In *Reports of the 6th International Congress on Acoustics* (Vol. 76; p. 10).

Smith, R., Braden, B. B., Chen, K., Ponce, F. A., Lane, R. D., & Baxter, L. C. (2014). The neural basis of attaining conscious awareness of sad mood. *Brain Imaging and Behavior*, 1-14.

Stein, J. L., Wiedholz, L. M., Bassett, D. S., Weinberger, D. R., Zink, C. F., Mattay, V. S., & Meyer-Lindenberg, A. (2007). A validated network of effective amygdala connectivity. *Neuroimage, 36*(3), 736-745.

Terhardt, E. (1977). The two-component theory of musical consonance. In E. F. Evans, & J. P. Wilson, *Psychophysics and Physiology of Hearing* (pp. 381-390). London: Academic Press.

Tervaniemi, M., Winkler, I., & Naatanen, R. (1997). Pre-attentive categorization of sounds by timbre as revealed by event-related potentials. *NeuroReport, 8*(11), 2571-2574.

Trainor, L. J., & Trehub, S. E. (1993a). Musical context effects in infants and adults: Key distance. *Journal of Experimental Psychology: Human Perception and Performance, 19*(3), 615.

Trainor, L. J., & Trehub, S. E. (1993b). What mediates infants' and adults' superior processing of the major over the augmented triad? *Music Perception*, 185-196.

van de Geer, J. P., Levelt, W. J., & Plomp, R. (1962). The connotation of musical consonance. *Acta Psychologia, 20*, 308-319.

Venkatraman, V., Rosati, A., Taren, A., & Huettel, S. (2009). Resolving response, decision, and strategic control: Evidence for a functional topography in dorsomedial prefrontal cortex. *The Journal of Neuroscience, 29*(42), 13158-13164.

von Helmholtz, H. (1877/1930). *On the Sensations of Tone as a Physiological Basis for the Theory of Music, translated by Alexander J. Ellis.* London: Longmans, Green, and Company.

Vossen, A. (2010). Shivers down the spine: Music as a pre-linguistic mnemonic device? *The Maastricht Journal of Liberal Arts, 2*, 3-16.

Vytal, K., & Hamann, S. (2010). Neuroimaging support for discrete neural correlates of basic emotions: A voxel-based meta-analysis. *Journal of Cognitive Neuroscience, 22*(12), 2864-2885.

Wetzel, W., Ohl, F. W., & Scheich, H. (2008). Global versus local processing of frequency-modulated tones in gerbils: An animal model of lateralized auditory cortex functions. *Proceedings of the National Academy of Sciences, 105*(18), 6753-6758.

Wicker, B., Keysers, C., Plailly, J., Royet, J. P., Gallese, V., & Rizzolatti, G. (2003). Both of us disgusted in my insula. *Neuron, 40*(3), 655-664.

Wright, P., He, G., Shapira, N. A., Goodman, W. K., & Liu, Y. (2004). Disgust and the insula: fMRI responses to pictures of mutilation and contamination. *Neuroreport, 15*(15), 2347-2351.

Zatorre, R. J., Belin, P., & Penhune, V. G. (2002). Structure and function of auditory cortex: Music and speech. *TRENDS in Cognitive Neuroscience, 6*(1), 37-46.

Zeki, S., & Romaya, J. P. (2008). Neural correlates of hate. *PLoS One, 3*(10), e3556.

Zentner, M., Grandjean, D., & Scherer, K. R. (2008). Emotions evoked by the sound of music: Characterization, classification, and measurement. *Emotion, 8*(4), 494-52.