In Search of a Soulful Sound: A Pianist's Exploration of Sonic Preferences

By

Woo Lee

A written project submitted in partial fulfillment
of the requirements for the degree of

Doctor of Musical Arts
(Piano Performance)

at the
UNIVERSITY OF WISCONSIN-MADISON
2023

Date of final oral examination: 12/08/2021

The written project is approved by the following members of the Final Oral Committee:

Christopher Taylor, Professor of Piano

Martha Fischer, Professor of Piano and Collaborative Piano

Margaret Butler, Associate Professor of Musicology

Brian Hyer, Professor of Music Theory

Table of Contents

IN SEARCH OF A SOULFUL SOUND:
A Pianist's Exploration of Sonic Preferences
Woo Lee
Doctor of Musical Arts: Piano Performance

ABSTRACT

The Doctoral Performance and Research submitted by Woo Lee, under the direction of Christopher Taylor at the University of Wisconsin-Madison, in the fulfillment of the requirements for the degree Doctor of Musical Arts consists of the following:

I. Written Project:

"In Search of a Soulful Sound: A Pianist's Exploration of Sonic Preferences"

This 44-page written Project is a research study aiming to understand the biological origins of sound preferences and explore their influences throughout human development. The study focuses on several topics: pitch, timbre, memory, noise, and learning. This research includes a survey conducted regarding sound preferences. Survey results revealed a negative correlation between sound preference and increases in timbral complexity and pitch instability.

II. Solo Recital, 12/14/2020, Remote
Pastorale for Organ, BWV590 - J.S. Bach/arr. Lipatti
Suite #1 for Keyboard, HWV434 - G. Handel
Sonata, Hob.XVI/28#43 - F.J. Haydn
Sonata, Op.28#15 - L.v. Beethoven
Etude-Tableaux, Op.39#8 - S. Rachmaninoff
Concert Etude, Op.40#6 - N. Kapustin

III. Solo Recital, 11/16/2021, Collins Hall
Kreisleriana, Op.6 - R. Schumann
Drei Klavierstucke, D.946 - F. Schubert

IV. Chamber Recital, 4/24/2022, Collins Hall
Violin Sonata SO.622#1 - L. Ornstein
Horn Trio, Op.40 – J. Brahms

V. Concerto Recital, 12/18/2022 Collins Hall
Piano Concerto, Op.20 – A. Scriabin

VI. Lecture-Recital, 5/13/2023 Collins Hall
"In Search of a Soulful Sound: A Pianistic Exploration of Color and Sounds"
An exploration of the nature of sounds and colors. Also includes performances of "Sheep May Safely Graze" arr. Egon Petri, Etude Op.25#7 by F. Chopin, Etude Op.25#7 arr. Glazunov, "Reflets dans l'eau" by C. Debussy, and "Les jeux d'eaux à la Villa d'Este" by F. Liszt.

VII. Final Solo Recital, 12/3/2023 Collins Hall

Impromptus, Op.90#1&3 - F. Schubert
Impromptu Op.29#1 - F. Chopin
Impromptu Op. 51#3, - F. Chopin
Impromptus, Op.5#1,2,3,&4 - J. Sibelius
Improvisations, FP.63#1,2,&3 - F. Poulenc
Improvisation, FP.176#15 - F. Poulenc
Impromptus, Op.68#1&2 - L. Liebermann
Impromptu, Op.66#2 - N. Kapustin

Introduction

Sounds are a crucial natural phenomenon that help humans navigate their surroundings, providing us with essential information about how to respond to the external world. Our ears can convert multiple streams of aural stimuli into electrical signals, allowing our brains to differentiate between sounds with staggering speed and precision. While the ear's primary function is to allow successful interactions with the outer world, humans have also come to enjoy sounds regardless of their immediate utility. People are subconsciously drawn towards certain sounds and repelled by others, and this is the foundation upon which all music is built.

Despite being intangible, invisible, and perhaps even "inessential," providing no nourishment to our bodies, it is indisputable that music is highly valued in societies and cultures across the world. The abstract and immaterial nature of music makes it a highly elusive topic for academic study; neither its influences nor its origins can be pinpointed without a generous dose of speculation. An early example of such speculation was penned by Charles Darwin, who proposed that "the habit of uttering musical sounds was first developed, as a means of courtship."[1] Similarly, regarding the question of why music has become so important to society, Ellen Dissanayake theorizes that the human inclination to predict future outcomes and gather social support made music an ideal communal coping mechanism for anxieties.[2]

It is undeniable that music fulfills important societal, emotional, recreational, and cultural roles, and these facts have propelled research directed at understanding its effects on individuals

---

[1] Charles Darwin, *Expression of the Emotions in Man and Animals* (London: John Murray, 1872), 87.

[2] Stephen Malloch, Trevarthen Colwyn, and Ellen Dissanayake, *Communicative Musicality: Exploring the Basis of Human Companionship* (New York: Oxford University Press, 2009), 24.

and communities. However, most studies evade the question of why we prefer certain sounds over others, and what we tend to prefer. Sounds are the building blocks of music and can conveniently be isolated and analyzed with the right tools and techniques; this is the focus of acoustic science and sound engineering, among many other disciplines. But while knowledge abounds in the fields of physics, biology, and psychology, all areas which are needed to understand musical perception, there is still more to be learned about the roots of human sonic preferences. Few sources provide consolidated information about this wide topic.

I believe the reason for the scarcity of research in this specific area is twofold. First, music perception lies at the intersection of several academic fields, many of which are not closely related, and this creates challenges for multidisciplinary work. For example, university professors in music performance are typically master practitioners who focus their efforts on preparing students for performing careers. Professors in the other academic fields can be knowledgeable in music, but they often lack the insights of a professional musician who has invested thousands of hours into the craft. Secondly, the realm of musical expression and sound perception is notoriously complex, context-dependent, and subjective; it fails to be understood because the human mind is difficult to understand. These complexities and contingencies, though inconvenient for analysts, are indeed what makes music and sounds an invaluable gateway into the human experience. The benefits of a comprehensive understanding of sounds, music, and ultimately ourselves, surely outweigh the challenges that present themselves during the pursuit of these many mysteries.

Purpose and Approach

Arguably, the persuasive power of music has less to do with the physical composition of sound waves themselves, and more to do with the value people ascribe to them. Even before modern scientific instruments made the detailed examinations of sounds and neurons possible, musicians were sensitive to something more enigmatic than the immediately obvious qualities of sound. Johann Sebastian Bach's first biographer, Johann Nikolaus Forkel, noted the following in *Johann Sebastian Bach: His Life, Art, and Work* (1802):

> Bach preferred the Clavichord to the Harpsichord, which, though susceptible of great variety of tone, seemed to him lacking in soul. The Pianoforte was still in its infancy and too coarse. Both for practice and intimate use he regarded the Clavichord as the best instrument and preferred to express on it his finest thoughts.[3]

Forkel focused particularly on two concepts: "tone" and "soul." "Tone" is an ambiguous term, though it seems to relate both to the terms "pitch" and "timbre." "Soul" is even hazier to define, yet it was apparently the decisive factor for J.S. Bach when comparing the clavichord with the harpsichord. What makes a sound "soulful"?

My intent in the following pages is to provide plausible responses for the following questions:

1. What influences sound preferences throughout human development?
2. What kinds of sounds do people prefer?

Research on sound perception can and should be drawn from exact sciences, but there is an inevitable degree of speculation that must be acknowledged; this essay is not intended to be

---

[3] Johann Nikolaus Forkel, *Johann Sebastian Bach: His Life, Art, and Work* (1802), trans. Charles Sanford Terry (London: Constable and Company, 1920), 58-59.

comprehensive, nor does it intend to put forth an objective framework of sound preferences. Such a framework, if it exists at all, would require many additional research efforts and diverse perspectives. The purpose of my thesis is to summarize and explore the biological mechanisms undergirding our perception of sound, and to present relevant studies in cognition that will help elucidate the complex matter of auditory preferences. Although I intend to conduct a rigorous investigation of the subject at hand, this research is presented from the perspective of a trained musician. I include the word "pianist" in the title to specify that my conclusions and overall approach may drastically differ from those of a cognitive or behavioral scientist. Lastly, due to the scope of this thesis, I mostly explore single sounds as opposed to multiple consecutive or simultaneous sounds (like those often heard in music). Numerous tones can potentially be perceived in relation to a musical syntax, and syntax is largely generated and influenced by culture. The impact of culture on aural preferences is a broad and intricate topic that exceeds the parameters of the present research but warrants further discussion by musicians, scientists, historians, and other experts.

While much of this research aims to provide a foundation for future academic endeavors, I believe it is both possible and fruitful to test my premises in the present. My research also includes a study designed to evaluate the sound preferences of students at the University of Wisconsin-Madison. This survey, which asks students to listen to specific sound samples and indicate their strength of attraction to each sound, is meant to support my hypotheses about the origins of and basis for our aural inclinations.

Multiple factors contribute to the formation of preferences, and a complex issue is best handled with a wide approach. The early sections will present preliminary information about our hearing apparatus, and later sections will combine the findings of multiple neurological and behavioral studies to piece together the origin and machinations of human preferential development. Following these explorations, I will present the survey, discuss its experimental design, organize the collected data, and form conclusions about the emergent patterns. At the conclusion, I will elaborate on the "soulful sound framework," synthesizing previous sections and explaining its relevance to musical aesthetics.

Hearing

To establish a foundation for the following discussion of sound preferences, some preliminary information regarding sounds, the ears, and the brain must be clarified. Sounds are primarily air pressure waves that may be broken into multiple individual sine waves. Sine waves themselves are extremely simple, oscillating at a predictable rate; however, typical everyday sounds are the result of many distinct sine waves which combine additively, creating a complex array of frequencies. What we hear as a "pitch" is the fundamental frequency, which is the lowest and usually the loudest component in a sound wave. Other frequencies present in the sound wave contribute to the "timbre" of a sound, which might be considered akin to its "taste."[4] The "volume" of a sound is determined by the wave's amplitude, which (in simple terms) is the amount by which individual air particles move away from their resting points.[5]

Here I will summarize the basic components of hearing, in order to establish a baseline regarding how perception begins. Sounds are collected by the ear before being transmitted to the

---

[4] Hermann von Helmholtz, *On the Sensations of Tone*, trans. Alexander Ellis (New York: Dover Publications, 1954), 56-58.

[5] Helmholtz, *On the Sensations of Tone,* 10.

brain for processing and storage. Generally, physiologists divide the ear into the outer, middle, and inner components. Sounds travel through the auditory canal of our outer ear before proceeding to the middle ear's eardrum (also known as the tympanic membrane), which they cause to vibrate. Connected to the eardrum are three small bones called the ossicles (the malleus, incus, and stapes). These attach in turn to the cochlea, which is a spiral tube containing fluid. The cochlea is separated into an upper and lower portion by the basilar membrane, and the two levels respond (respectively) to low-pitched frequencies and high-pitched frequencies. Along the basilar membrane are frequency-specific hair cells that resonate upon contact with the fluid wave created by the ossicles. The hairs then produce auditory information, sent via the cochlear nerve into the central auditory system. This system includes various higher order auditory structures and ultimately ends at the auditory cortex.[6]

The beginning of auditory perception is the identification of the sound object. Sounds from separate sources do not enter the central auditory processing system as distinct entities; individual waves are simply added cumulatively to each other before entering the auditory system. The complexity of the entire system is what allows us to process the endless varieties of vibrations that our ears encounter. The brain's ability to separate out the components of composite sounds, the phenomenon of perceptual auditory object formation, is called stream segregation, or "streaming." According to Daniel Pressnitzer, this is what allows humans to

---

[6] Hinrich Staecker and Jennifer Thompson, "Central Auditory System, Anatomy," in *Encyclopedia of Otalaryngology, Head and Neck Surgery* (Berlin: Springer-Verlag Heidelberg, 2013): 376–383.

perceptually organize their acoustic environments and pinpoint sound sources such as "a musical instrument within an orchestra."[7]

The separate streams of sound that we hear appear to be handled by separate groups of neurons before they are stored in memory. Pressnitzer explains this in more detail:

> The general form of the neural correlates… can be described as "grouping by coactivation": Sounds that activate the same or largely overlapping populations of neurons are perceived as forming a single stream, whereas sounds that activate different neuronal populations are perceived as separate streams. [8]

He goes on to mention that, in a scenario where multiple tones of varying frequencies are heard, tones with similar frequencies are categorized in a single stream by neural correlates, while tones that are dissimilar in frequency are heard in separate streams.[9]

Pitch

Of the many qualities of a sound, pitch is a large factor influencing successful perception and storage in memory. On this topic, Andrew Oxenham writes that "there are a number of acoustic cues that influence the perceptual organization of simultaneous sounds, such as onset and offset asynchrony (whether components start and stop at the same time) and harmonicity (whether components share the same fundamental frequency or F0)."[10] Oxenham mentions that

---

[7] Daniel Pressnitzer, Mark Sayles, Christophe Micheyl, and Ian M. Winter, "Perceptual Organization of Sound Begins in the Auditory Periphery," *Current Biology* 18.15 (2008): 1124–1128.

[8] Pressnitzer et al., "Perceptual Organization Sound Begins," 1124.

[9] Ibid.

[10] Andrew J. Oxenham, "Pitch Perception and Auditory Stream Segregation: Implications for Hearing Loss and Cochlear Implants," *Trends in Amplification* 12.4 (2008): 316.

pitches seem to activate parts of the temporal lobe,[11] which is associated with hearing, speech perception, and certain aspects of memory, and Joshua McDermott proposes that the ability to detect pitch changes in two or more different tones in sequence is central to musical enjoyment.[12] The latter conclusion is supported by an in-depth study of congenital amusia (tone-deafness), in which a tone-deaf subject was asked to indicate whether a pitch change occurred in a five-note string of otherwise identical pitches. The study showed that she could "barely detect a rising pitch change as large as two semitones," and that "she is at chance for one-semitone changes."[13] McDermott also remarks that, in a study by Isabelle Peretz et al., a subject who was once a music enthusiast reported disliking music after suffering brain lesions in the left temporal and right frontal lobes; after his injuries, he was "found to be insensitive to scale violations ("sour" notes, to which untrained Western listeners are acutely sensitive)."[14]

This is not to say that unpitched sounds are inferior to pitched sounds from a neurological perspective; percussive instruments, for example, are a welcome component in many musical genres. Instead, pitch (along with other aspects of sound like timbre or rhythm) is a key characteristic that our brains understand and use to define sounds during memory acquisition. Understanding how our neurology has adapted to receive pitched sounds more easily can provide us with necessary context for discussing the purpose of our specialized hearing, as well as the role memory plays in guiding our inclinations.

---

[11] Josh H. McDermott and Andrew J. Oxenham, "Music Perception, Pitch, and the Auditory System," *Current Opinion in Neurobiology* 18.4 (2008): 456.

[12] McDermott, "Music Perception, Pitch, and the Auditory System," 456.

[13] Isabelle Peretz, Julie Ayotte, Robert J. Zatorre, Jacques Mehler, Pierre Ahad, Virginia B. Penhune, and Benoît Jutras, "Congenital Amusia: A Disorder of Fine-Grained Pitch Discrimination," *Neuron* 33 (2002): 188.

[14] Joshua McDermott and Mark Hauser, "The Origins of Music: Innateness, Uniqueness, and Evolution. *Music Perception* 23.1 (2005): 49.

In music, pitches (and silence) are the building blocks necessary for the expression of ideas, emotions, and narratives. Melodic contours convey shape and structure in a way that the brain can easily recognize. A study in 1980 presented new melodic material to non-musician adults, and the results showed that the subjects were "likely to remember little more than its contour."[15] An exception is that, with well-known melodies, most adults can perceive relatively subtle alterations in interval content and tuning.[16] Contour, the quality of pitch movement across time, is a guide rail that the act of musical prediction depends on. The significance of prediction in music will be discussed in a later section, but it suffices here to say that the cycle of prediction and satisfaction is a key component of musical enjoyment — regarding the role of prediction in melodic perception, Stephen Malloch writes that "the human ability to sense the shape of a melody within the body is intrinsic to our enjoyment of music as human communication."[17] Modern developmental studies confirm that infants demonstrate sensitivity for melodic phrase structure, attending to the rhyming vowels at the ends of lines, and by five months the infant can vocalize a matching vowel in synchrony with the mother.[18] All of these abilities highlight the innate human sensitivity to pitches, gestures, and shapes, and sensing these elements of sounds plays a large role in experiencing musical enjoyment.

---

[15] Sandra Trehub, Glenn Schellenberg, and David Hill, "The Origins of Music Perception and Cognition: A Developmental Perspective*,"* In *Perception and Cognition of Music* ed. Irène Deliège and John A. Sloboda (New York: Psychology Press, 2015), 106.

[16] Ibid.

[17] Stephen Malloch and Colwyn Trevarthen, "The Human Nature of Music," *Frontiers in Psychology* 9 (2018): 7.

[18] Malloch, "Human Nature," 7.

Timbre

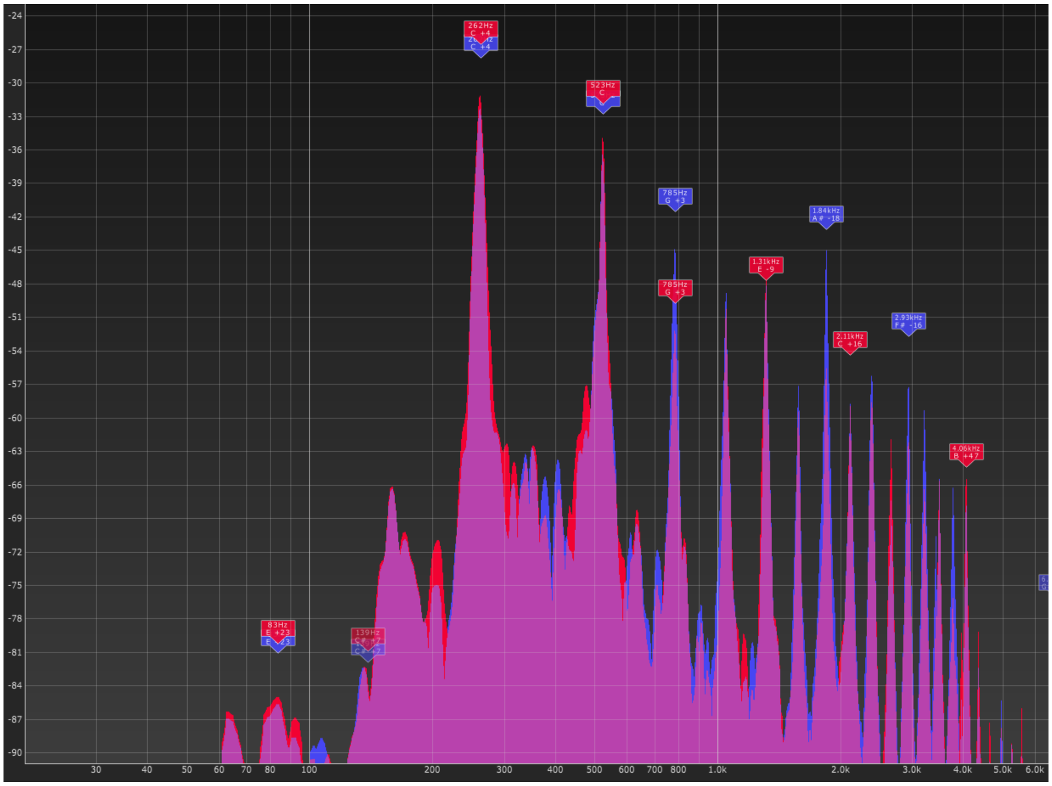
Timbre is one of the most complex qualities of sounds, since one cannot begin to describe it in terms of a single acoustic parameter. While volume, for example, can be defined simply by amplitude, and pitch involves the perception of the fundamental partial (explained in more detail below), timbre depends on the total harmonic spectrum of a sound and manifests in a seemingly infinite number of sonic variations in the natural world.

The *Encyclopaedia Britannica* defines musical timbre as “the characteristic tone color of an instrument or voice, arising from reinforcement by individual singers or instruments of different harmonics, or overtones of a fundamental pitch.”[19] Somewhat analogous to the way that the taste of foods can be broken down into numerous chemical components, timbre depends on a multiplicity of partials. As discussed in a previous section, naturally occurring sound waves are combinations of individual sine waves, and the lowest and often loudest frequency determining the basic pitch is the “fundamental.” A partial is “one of the component vibrations at a particular frequency in a complex mixture,”[20] and certain partials that are exact multiples of the fundamental frequency are also called “overtones”.[21] The overtone above the first partial (which is another name for the fundamental) is called the second partial, and so on. This phenomenon can be observed with a spectrograph, a scientific instrument that can translate sound waves into a visual diagram (a spectrogram). An example spectrogram of a grand piano playing C4 is shown below, with frequency across the x-axis and amplitude on the y-axis. This is a snapshot of the sound wave near its beginning.

---

[19] *Britannica.com,* “Timbre” (accessed October 14, 2023).
[20] Murray Campbell, “Partial,” *Grove Music Online* (accessed October 14, 2023).
[21] *Britannica.com,* “Overtone” (accessed October 14, 2023).

Figure 1: Spectrogram of a Grand Piano Playing C4 (fundamental at 262 Hz).



For musical tones, the prominence of certain overtones greatly influences the color of the sound. For comparison, a spectrogram of a cello playing A3 has also been provided. The cello’s partials in the upper frequency ranges are particularly active, contrasting with the frequencies observed in the piano’s spectral profile.

Figure 2: Spectrogram of a cello playing A3 (fundamental at 221 Hz)

Robert Cogan addresses the impact of partials on timbre in *Sonic Design*, focusing particularly on Hermann von Helmholtz, a key figure in the history of psychoacoustics. In 1863, Helmholtz hypothesized that "differences of tone color arise principally from the combination of different partial tones with different intensities," which Cogan states was the "beginning of tone-color analysis."[22] Modern spectrographic technology confirms the ways in which the number of partials and their intensities (amplitudes) are responsible for timbre.

Helmholtz and his successors have made various attempts to describe the human response to tone color. Helmholtz's framework divided sounds into certain classes based on which specific partials were observed:

---

[22] Robert Cogan, *Sonic Design: The Nature of Sound and Music* (New Jersey: Prentice Hall International, 1976), 329.

> First class: Simple sine tones, like those of a tuning fork mounted on a resonator, and wide-stopped organ pipes; they have a soft, pleasant sound, free from all roughness, but wanting in power and dull at low pitches.
>
> Second class: Notes accompanied by a moderately loud series of the lower partial tones up to about the sixth partial, like those of the middle register of the piano, open organ pipes, and the softer tones of the human voice and of the French horn; these are more harmonious and musical, rich and splendid compared with simple tones, yet at the same time sweet and soft if the higher partials are absent.
>
> Third class: Those of narrow stopped organ pipes and the clarinet, that give only the odd-numbered partials, producing a hollow, even nasal quality of tone. When the fundamental tone predominates, the quality is rich; when the fundamental is weak, the quality is poor.
>
> Fourth class: Those notes of bowed instruments, of most reed pipes of the organ, and of the oboe, the bassoon, the harmonium, and the human voice (certain vowels) whose partials above the sixth are prominent, the quality of tone being cutting and even rough.[23]

James Jeans proposed that individual partials, particularly at the lower perceptible levels, have somewhat consistent effects on sound:

> The second partial adds clearness and brilliance.
>
> The third partial adds brilliance but also a certain hollow, throaty, or nasal quality.
>
> The fourth partial adds more brilliance, and even shrillness.
>
> The fifth partial adds a rich, somewhat hornlike quality to the tone.
>
> The sixth partial adds a delicate shrillness of nasal quality.[24]

Cogan points out that these provide a starting point for tone-color analysis, particularly when analyzing instruments whose spectral profiles change according to dynamic, register, and technique. About the clarinet, Cogan writes:

> Throughout the clarinet range, there is a predominance of fundamental and odd-numbered partials…. The louder a tone, the more extended is its series of upper

---

[23] Cogan, *Sonic Design,* 329.
[24] Ibid.

> partials and the greater their relative intensity. The elimination of even-numbered partials may account for the "hollowness" of the clarinet sound. The presence in quantity of the fifth, seventh, and ninth partials accounts for the reediness, brilliance, and harshness of certain loud clarinet colors.[25]

Several studies have tested the brain's recognition of timbre, though "timbre perception remains less precisely understood than that of the other auditory attributes, both at the behavioral and neurophysiological levels."[26] One study assessed timbral memory by presenting original instrument sounds and subsequent "transformed" sounds with altered spectral characteristics, and the results showed that timbre is in fact relevant to successful memory storage. Other studies testing the role of timbre in memory have shown that selectively altering the spectral characteristics of a known sound noticeably impairs recognition.[27]

---

[25] Cogan, "Sonic Design," 355.

[26] Anne Caclin, Elvira Brattico, Mari Tervaniemi, Risto Näätänen, Dominique Morlet, Marie-Hélène Giard, and Stephen McAdams, "Separate Neural Processing of Timbre Dimensions in Auditory Sensory Memory." *Journal of Cognitive Neuroscience* 18.12 (2006): 1959.

[27] Lutz Jäncke, "Music, Memory and Emotion," *Journal of Biology* 7.21 (2008): 1.

Memory

Temporality is inseparable from sounds; wave patterns cannot exist without the passage of time, and we cannot process wave patterns without memory. Neuropsychologist Lutz Jäncke explains that "musical sounds, like all auditory signals, unfold over time. It is therefore necessary for the auditory system to integrate the sequentially ordered sounds into a coherent musical perception."[28] Auditory sequencing in the brain "can be considered a mechanism of working memory,"[29] which in Esther Heerema's formulation is "the ability not only to remember information for a period of time but also to use, manipulate and apply it, perhaps while also accessing other stored pieces of information."[30] Sounds are held in the working memory in a way that resembles how we visualize "an arithmetic problem without paper."[31]

At the core of our inclinations are memories and experiences that serve as the basis of attraction and repulsion; as is the case with our culinary appetites, our ears are drawn towards sounds that have been associated with positive experiences in the past. Auditory storage, like our memory for other senses like sight and smell, is not restricted to purely sensory experiences; emotional responses are connected to our memories of sounds. The activation of emotions and heightening of the attention is largely a motivational mechanism of the limbic system, which is "involved in processing of emotions and in controlling memory."[32] Regarding the links between music and emotions, Jäncke writes that "positive emotions and high arousal levels that are

---

[28] Jäncke, "Music, Memory, and Emotion," 1.
[29] Ibid.
[30] Esther Heerema, "How Working Memory is Affected by Alzheimer's Disease," *verywellhealth.com* (accessed October 14, 2023).
[31] Nelson Cowan, "What are the Differences Between Long-Term, Short-Term, and Working Memory?" *Progress in Brain Research* 169 (2008): 326.
[32] Jäncke, "Music, Memory, and Emotion," 1.

associated with specific events act as a memory enhancer for these particular events. In the context of associative memory models, this memory-enhancing effect… can be explained as a strengthening of the associations between the memories due to strong emotions and to arousal."[33] The body remembers not only modality-specific auditory data, but also the feelings of emotional "valence" associated with them. In a web of memory connections, emotions play a role in reinforcing connections and recalling specific events.

A study titled "Unforgettable Film Music: The Role of Emotion in Episodic Long-Term Memory for Music" investigates the effects of emotion on musical memory, asking 24 participants to rate selections of film music across three sessions based on arousal, valence, and felt emotional intensity. In the first session with subjects, 40 musical selections were played. The second session consisted of re-listening to the previous session's selections. The third session consisted of 40 old and 40 new musical excerpts, with the same rating tasks assigned as in the first session. "Arousal" and "valence" were defined by the study as follows:

> Arousal: The excitation level elicited by the music (ranging from very relaxing to very exciting).
>
> Valence: The emotional value (on a continuum from negative to positive, unpleasant to pleasant) elicited by a musical stimulus.[34]

The results of the study showed that "very positive valence ratings seem to be associated with better memory performance of music in a recognition task."[35] This seems to complement the findings by another study which concluded that happy and peaceful music induced the

---

[33] Jäncke, "Music, Memory, and Emotion," 4.

[34] Susann Eschrich, Thomas F. Münte, and Eckart O. Altenmüller, "Unforgettable Film Music: The Role of Emotion in Episodic Long-Term Memory for Music," *BMC Neuroscience* 9.1 (2008): 2.

[35] Eschrich, "Unforgettable Film Music," 5.

strongest emotions,[36] although other studies also that show that strongly negative stimuli can have a similar effect on sound recognition.[37] This makes sense considering the purpose of emotions: They guide us towards positive stimuli and away from negative stimuli, and neural pathways are strengthened through repetition.

A challenging question regarding perception is how familiar sounds are perceptually grouped together, and whether they can conjure similar emotions. Do all high-frequency bird calls inspire tranquility, or is there a certain point when a bird call becomes frightening? Intuitively, some subjective responses to certain sounds are common among most people, and both context and intensity can impact our perceptions of those sounds. Deep low-frequency storm tremors can inspire panic when one is outside, but some enjoy distant rumblings of storm clouds from the comfort of their homes (in this case, elevated levels of situational reassurance and safety are likely to serve as emotional enhancers). Likewise, high-frequency bird calls heard in the morning can trigger relaxation and cheerfulness, while frenzied screeches of a frightened animal can trigger alertness. Sound perception is heavily affected by setting; Andrew Oxenham notes that "beyond these acoustic, or bottom up, cues, there are many top-down influences on perceptual organization, including attention, expectations, and prior exposure, all of which can also play an important role."[38] The implication is that the mind is attentive to contextual information as well as prior expectations, which are a function of memory.

A potential reason for this flexibility of sound perception is that short-term memories are subject to substantial feature decay over time. István Winkler asserts that "it has been found in many studies that subjects can only tell the difference between two closely similar sounds if the

---

[36] Eschrich, "Unforgettable Film Music," 2.
[37] Ibid., 5.
[38] Andrew Oxenham, "Pitch Perception and Auditory Stream Segregation," 316.

sounds to be compared are presented within ca. 10 s," and also that interrupting two similar sounds with dissimilar sounds "deteriorates performance" when the compared sounds themselves are similar.[39] Long term memories are generally blurry; Winkler points out that sensory stores have a "cruder resolution" when the sensory trace of the initial sound is gone.[40] It might be that the observed lower sensory precision of memories allows for some fluidity regarding memory recollection and emotional responses.

A possible exception to Winkler's assertion is speech information, which persists longer than "the 30 s or so that has been the presumed duration of auditory sensory memory."[41] One of Winkler's experiments tested the effects of vocal information on memory recollection, by playing instances of two different words with the same meaning, at a delayed interval. The result showed that the second word was consistently recognized faster and more accurately when the voice remained the same across both words. It could be that spoken words are easier to remember for longer because they are rich in codifiable information such as semantics, syntax, and phonetics, and that humans are highly specialized for communication and speech comprehension. The brain does not process non-verbal sounds with the same degree of specificity, although it does recognize musical syntax in organized pitches, as discussed earlier. This might be one reason why musical sounds can evoke a wide and relatively unspecific range of emotional responses, despite being easily codifiable in memory when arranged according to a known system.

---

[39] István Winkler and Nelson Cowan, "From Sensory to Long-Term Memory," *Experimental Psychology* 52.1 (2005): 1–2.

[40] Winkler, "From Sensory to Long-Term Memory," 4.

[41] Ibid.

Noise

Aversions can reveal just as much about our predilections as our proclivities. One topic relevant to the questions explored here concerns the definition of the word "noise," which Merriam-Webster defines as a sound that "lacks an agreeable quality or is noticeably unpleasant or loud."[42] This term is clearly quite loose; a team of researchers (Liu et al.) wrote that "although researchers have investigated the ability of people to listen, analyze, and distinguish sound, the concept of noise has not been clearly articulated from a human perspective."[43] The team sought to examine what people perceive as noise through qualitative interviews, and they discovered that noise is 1) context and environment-dependent, 2) related to the objective quality of the sound itself, and 3) influenced by subjective elements such as the listener's age, occupation, and preferences.

Liu et al. observed that noise is frequently associated with a "roughness" and "sharpness"[44] in the sound. Josh McDermott clarifies some of these terms in *Auditory Preferences and Aesthetics: Music, Voices, and Everyday Sounds* (2012)*:*

> "Sharpness describes the proportion of energy at high frequencies, found to be less pleasant. Frequencies in the range of 2-4 kHz contribute the most annoyingness."[45]
>
> "Roughness is the perceptual correlate of fluctuations in energy (intensity) that occur over time, analogous to the fluctuations in surface depth that determine the roughness of an object to the touch. Fluctuation at rates between ~20-200 Hz [referring to volume

---

[42] *Merriam-Webster.com*, "noise" (accessed October 11, 2023).
[43] Fangfang Liu, Shan Jiang, Jian Kang, Yue Wu, Da Yang, Qi Meng, and Chaowei Wang. "On the Definition of Noise," *Humanities and Social Sciences Communications* 9.1 (2022): 1.
[44] Liu, "On the Definition of Noise," 8.
[45] McDermott, Josh H. "Auditory Preferences and Aesthetics: Music, Voices, and Everyday Sounds," in *Neuroscience of Preference and Choice*, edited by Raymond Dolan and Tali Sharot, Cambridge: Academic Press (2012): 228–229.

> oscillations] are those that determine roughness; any lower, and the fluctuations can be heard individually rather than contributing to a sound's timbre."[46]

Concerning sharpness, he notes that sound waves in the 2-4 kHz range are amplified by "as much as 30 dB" when entering the ear canal, as proposed by the research of Henoch et al., making sounds in this range the "most likely to damage the ear." [47] These sounds become significantly less bothersome at lower volume levels.[48] Louder sounds present a threat to the auditory system,[49] which is reflected in the National Institute of Occupational Safety and Health's recommended exposure limits. NIOSH specifies a limit of 85 dB (comparable to a noisy restaurant) for a span of 8 hours, and recommended exposure time is halved each time the volume increases by 3 dB; longer or louder exposure constitutes a risk to hearing.[50]

Aversion towards roughness is more perplexing, McDermott explains, "as it does not obviously pose any danger to the auditory system."[51] His definition of roughness is derived from Ernst Terhardt, who explains that "the physical parameter which almost exclusively determines roughness is the relative amplitude fluctuation," and that "beyond about 20Hz [or 20 volume oscillations per second] fluctuations are perceived as an unpleasant, disturbing component which usually is called 'roughness,' 'raucousness,' or 'harshness'."[52] Conversely, a musical technique such as vocal vibrato utilizes both volume and pitch fluctuations; according to Carl Seashore,

---

[46] McDermott, "Auditory Preferences and Aesthetics," 229-230.

[47] Miriam A. Henoch and Kris Chesky, "Ear canal resonance as a risk factor in music-induced hearing loss," *Medical Problems of Performing Artists* 14.3 (1999): 103–106.

[48] Ibid., 106.

[49] Ibid., 106.

[50] Chuck Kardous, Christa L. Themann, Thais C. Morata, and Gregory Lotz, "Understanding Noise Exposure Limits: Occupational *vs.* General Environmental Noise," *Centers for Disease Control and Prevention* (blog accessed October 11, 2023).

[51] McDermott, "Auditory Preferences," 230.

[52] Ernst Terhardt, "On the Perception of Periodic Sound Fluctuations (Roughness)," *Acta Acustica United with Acustica* 30.4 (1974): 201.

vocal vibrato is "a pulsation of pitch, usually accompanied with synchronous pulsations of loudness and timbre" which gives tones "a pleasing flexibility, tenderness, and richness."[53] Concerning vibrato rates, Johan Sundberg conducted a study of ten professional opera singers and found that "the average vibrato rate found was 5.4 Hz for the male singers and 5.9 Hz for the female singers."[54] Sundberg provides some approximate guidelines, saying that "a rate slower than 5 undulations per second tends to sound unacceptably slow, and vibrato rates exceeding 8 undulations per second tend to sound nervous."[55] Although his statement is probably an informed personal observation rather than a scientific conclusion, it might provide some clues concerning why rapid volume changes are considered undesirable. I propose that most sounds we associate with extreme oscillation rates tend to signal danger, like the rumblings of an earthquake, or are similar to coping responses, like shivering in cold temperatures or trembling in fear. Ultimately, McDermott concedes that the reasons why roughness is perceived as unpleasant are still unclear.

Roughness can even be *pleasant* in the right context; compared with sharpness, roughness and our reaction to it seems highly dependent on cultural and contextual factors. For example, rough sounds enjoy frequent usage in rock and other similar genres, and the aforementioned use of vibrato has persisted throughout modern Western music history. One reason why vibrato is enjoyed, Sundberg suggests, might be because it "makes the singer's voice easier to discern against the background of a loud orchestral accompaniment."[56] Additionally, vocal vibrato expresses vocal ease and freedom, informing the listeners that "the singer is

---

[53] Carl Seashore, "The Psychology of Music. VI. The Vibrato: (1) What Is It?," *Music Educators Journal* 23.4 (1937): 30.

[54] Johan Sundberg, "Acoustic and Psychoacoustic Aspects of Vocal Vibrato," *Dept. for Speech, Music and Hearing Quarterly Progress and Status Report* (1994): 50.

[55] Ibid., 57.

[56] Ibid., 64.

solving a difficult vocal task without a struggle."[57] This research seems to reveal that humans generally enjoy a small amount of fluctuation in sounds, and this statement is supported by Seashore who claims that humans are accustomed to periodic pulsations, evident in natural behaviors such as laughter.[58]

Returning to the subject of ambiguity inherent in the term "noise," three questions arise:

1. How is noise most commonly defined?
2. What other general properties do noises possess?
3. Is noise universally disdained?

There are many meanings of the word "noise," varying according to context. It seems likely that a purely scientific account of the word will remain elusive, though various scholars have wrestled with the issue, beginning with Helmholtz in *On the Sensations of Tone* (1863). His definition seems to focus on sounds in nature, such as leaf-rustling or splashing; he states that "noise is accompanied by a rapid attenuation of different kinds of sensations of sound…. In all these cases we have rapid, irregular, but distinctly perceptible alternations of various kinds of sounds, which crop up fitfully."[59] Turning to "musical tone," Helmholtz writes:

> On the other hand, a musical tone strikes the ear as a perfectly undisturbed, uniform sound which remains unaltered as long as it exists, and it presents no alternation of various kinds of constituents. To this then corresponds a simple, regular kind of sensation, whereas in a noise many various sensations of musical tone are irregularly mixed up and as it were tumbled about in confusion.[60]

---

[57] Ibid., 65.

[58] Seashore, "The Psychology of Music," 31.

[59] Helmholtz, *On the Sensations of Tone*, 7.

[60] Helmholtz, *Sensations of Tone,* 8.

What Helmholtz appears to be referring to is that musical tones are uninterrupted, display regularity (he provides no clarification to the term "regularity", but we might assume he is referring to a waveform that contains repetitious patterns over a significant length of time), and create a simpler "sensation" (a changing sound, by contrast, complicates the auditory data traveling into the ear).

In contrast with Helmholtz, the Encyclopaedia Britannica offers a definition focused on noise's negative effects: Noise is "any undesired sound, either one that is intrinsically objectionable or one that interferes with other sounds that are being listened to."[61] This interpretation depends on the judgement or goals of the listener. A quiet conversation, for example, may be considered noise when it occurs during a concert of classical music, but it provides relaxed ambience in a coffee shop. In all, the Encyclopaedia Britannica's scope is significantly broader than Helmholtz's and avoids pointing to any specific sonic characteristics.

The term "noise" can also be used in a non-auditory framework, as well. In information theory and electronics, according to the Encyclopaedia Britannica, "noise refers to those random, unpredictable, and undesirable signals, or changes in signals, that mask the desired information content."[62] While the Encyclopedia Britannica authors here focus on electronic signals, traits such as "randomness" and "unpredictability" align with Helmholtz's observation that "irregularity" and "confusion" are qualities of noise.[63] This characterization makes sense from a biological perspective: The more regular or pattern-like a sound is, the more likely the brain is to recognize it. A "chaotic" or "unstable" sound is less likely to be remembered accurately. And on

---

[61] *Encyclopaedia Britannica.com*, "Overtone" (accessed October 14, 2023).

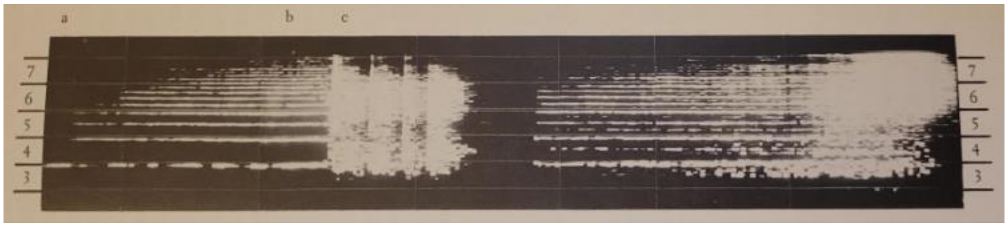
[62] *Encyclopaedia Britannica.com,* "noise" (accessed October 14, 2023).

[63] Helmholtz, *Sensations of Tone,* 8.

the extreme end, irregular or unpredictable sounds can be sources of potential danger that our senses do well to warn us about.

Other approaches to noise focus on timbre. White noise, for example, is defined in Merriam-Webster as “a heterogeneous mixture of sound waves extending over a wide frequency range.”[64] This is demonstrated clearly in the spectrogram below, made by Robert Cogan depicting a moment in Alban Berg’s *Wozzeck* (Act 3, Scene 2). Clear pitches at a and b (from pitched orchestral instruments, shown in the upper margin of the spectrogram) are seen in distinct bands roughly between the 4th and 7th octaves, indicating the presence of a strong fundamental pitch and accompanying partials. This contrasts with the spectral profile of the loud bass drum strokes at c, which spreads across the 4th and 7th octaves but does not contain the clear horizontal bands of a and b.[65]

Figure C: from *New Images of Musical Sound.*[66] This spectrogram displays octave registers on the y-axis (for example, C in the fourth octave would be C4) and time across the x-axis.



Given the intense and consistent spectral complexity of the bass drum across a wide range of frequencies, should it be considered noise? What about other percussive instruments that “lie in between” in harmonic richness and volume, such as concert triangles, congas, or

---

[64] *Merriam-Webster.com*, “White Noise Definition & Meaning” (accessed October 11, 2023).
[65] Cogan, *New Images of Musical Sound*, Cambridge: Harvard University Press, 1984), 94.
[66] Ibid., 94.

timpani? At what point do sounds become noise? Cogan states that "sounds that apparently lack pitch—often called *noise*— are actually complexes of many pitches…. The distinction between sounds in which a single pitch is clearly perceived and those in which none is perceived is extremely hazy." He asserts that "the difference between noise and tone is one of degree."[67]

Returning to the questions I posed earlier and summarizing the foregoing discussion, I propose:

1. Possible types of definitions for noise: There are multiple definitions, but two types seem to stand out. One is a definition that depends on whether or not the listener dislikes the sound or determines it to be unhelpful or distracting. Others anchor the meaning on features such as volume or timbre, but the lines between noise and sound remain blurred.
2. General properties of noise: Most descriptions claim that certain traits such as unpredictability and undesirability are traits of noise. Cogan suggests that to be categorized as such, the sound must contain an extreme density of pitches. Environmental context, subjective preference, volume, and to a certain extent, spectral profile, are characteristics that influence whether a sound is considered noise.
3. On the universal disdain for noise: Depending on one's definition, some noises can be enjoyed under the appropriate circumstances (such as sounds heard during loud rock concert or a bass drum playing in a symphony). However, common uses of the term are subjective and linked with negative connotations. From this perspective, "enjoyable noises" are not truly noises to begin with.

---

[67] Cogan, *Sonic Design*, 439.

A Brief Discussion of "Regularity" in Musical Aesthetics

Even though music is an art form where a wide variety of sounds are encouraged, some tones are considered more desirable than others. A musician's skill in tone organization and production influences whether the sounds they produce cause comfort or discomfort. Regarding vibrato, for example, Sundberg claims that "the more skilled the singer, the more regular the undulations."[68] Regularity and tone quality appear to occupy similar spaces in piano pedagogy; a lyrical melody in the right hand, for example, requires the pianist to play notes within a relatively consistent volume range for the line to register in the mind as a single voice. A skilled pianist has precise control over all sounds produced by the instrument, and many believe that piano sounds from an expert can seem "more beautiful" than others. This is evident in the musical philosophy of pianist Josef Lhevinne, who achieved worldwide recognition at the turn of the twentieth century. In *Basic Principles in Pianoforte Playing* (1924), he expressed the idea that musical beauty begins with a "ringing, singing tone." He contrasts the beautiful "ring" of controlled playing with that of "bangy" playing, saying "the reason why a number of people say they do not care for piano playing is that so many so-called performers upon the instrument treat it as though it were an anvil and go on hammering out musical horseshoes."[69] A beautiful sound is one in which "there is lightness, fineness, regularity of playing, but without weakness or uncertainty,"[70] while a harsh sound involves physical and sonic "jerkiness."[71] Smoothness and a certain level of predictability seem integral to Lhevinne's conception of beautiful piano sounds.

---

[68] Sundberg, "Acoustic and Psychoacoustic Aspects," 48.

[69] Josef Lhevinne, *Basic Principles of Pianoforte Playing* (1924) (New York: Dover Publications, 1972), 22.

[70] Ibid., 25.

[71] Ibid., 23.

Arnold Schoenberg's experiments with dissonance, starting perhaps as early as the String Quartet No. 2, Op. 10, led him eventually to the pinnacle of artistic regularity and orderliness: twelve-tone serialism. While many concertgoers may view Schoenberg's legacy as one of chaotic dissonance, serialism is clearly a system grounded in rigorous organization. Nicholas Ruwet writes that Schoenberg's serialism was a response to the disorderly state of tonality and form at the time, saying "Indeed, it was introduced to keep alive a certain unity, even a certain uniformity, after the collapse of the system of tonal relations in the musical language."[72] In a sense, Schoenberg's musical structuralism may have been the result of a "homesickness" for order, regarding which Henri Pousseur commented as follows:

> When we listen to the music of Schoenberg it becomes apparent immediately that the symmetries found within it . . . are the fruits of a harsh irony, or, at most, of bitter nostalgia. They are proof enough that its composer is still dreaming of what he considers a "golden age" (though it may have been guilt), while at the same time he is well aware that it has long since vanished into the past.[73]

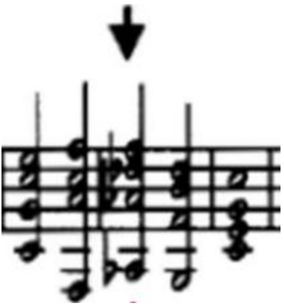
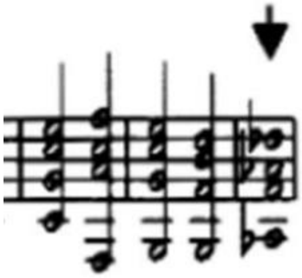
While the yearning for system and order can be observed in Western music history, this does not mean that the concept of regularity is inevitably superior to irregularity. For example, Western classical music employs irregularity and regularity in harmonic, rhythmic, and melodic structures to convey tension and release. Additionally, styles rooted in popular music such as jazz lean heavily on melodic, timbral, and rhythmic irregularity.

---

[72] Nicolas Ruwet, "Contradictions within the Serial Language," *Die Reihe* 6 (1964), 67.

[73] James Stiles, "The Decline of Serialism and the New Romanticism: Control and Chance in the New Music," *College Music Symposium* 19.1 (1979): 95.

Language, Prediction, and Purposeful Learning

Many thinkers intuitively feel that music is related to language; music communicates ideas, contains features such as rhythmic structure, syntax, and vocabulary, and has the potential to express emotions. On top of these characteristic similarities, science has increasingly supported the realization that music and language share common regions in the brain. One study titled "Musical Syntax is Processed in the Broca's Area (2001)" sought to discover whether musical syntax exists to the minds of listeners, and whether the brain processes musical syntax in the same area as language structures. The study measured event-related brain potentials (ERPs) in relation to harmonic incongruities, hypothesizing that ERPs in response to certain harmonically unexpected chords can reflect the existence of a musical type of syntax.[74] The researchers measured a particular ERP in non-musicians as they listened to Neapolitan chords within an excerpt following standard-practice tonality. Neapolitans were inserted in two different positions in a chorale-style context: One in the middle before a dominant cadence (D-flat major moving into G7), and one at the end after a dominant cadence (G7 moving into D-flat major). The excerpts used are shown in Examples 1 and 2, with arrows indicating the "third" and "fifth" positions in each excerpt where the Neapolitan was placed.

Example 1

Example 2

---

[74] Burkhard Maess, Stefan Koelsch, Thomas C. Gunter, and Angela D. Friederici, "Musical Syntax is Processed in Broca's Area: An MEG Study," *Nature Neuroscience* 4.5 (2001): 540.

The experimenters found that “the magnetic effect elicited by the Neapolitans was stronger at the fifth compared to the third position [in relation to the dominant chord], indicating that this effect reflects music syntactic processing. This effect was generated in both hemispheres in the *inferior pars opercularis*, known in the left hemisphere as Broca’s area (which is a region heavily involved in speech function.”[75] This study reveals not only that “these brain areas process considerably less domain-specific syntactic information than previously believed,” but also that “a strong relationship between the processing of language and music” exists from an anatomical perspective. Like with language, the brain uses syntax to predict and evaluate incoming musical sounds.

These advanced neurological systems fundamentally exist for the purpose of facilitating and enhancing predictive capabilities. Purposeful learning can be seen clearly in studies of newborns who have not been shaped yet by social dynamics and cultures. In a study titled “Of Human Bonding: Newborns Prefer their Mothers’ Voices (1980)”, a team of researchers sought to investigate the ability of newly born children to distinguish their mother’s voice from other female voices. The premise of the experiment was that “human responsiveness to sound begins in the third trimester of [pregnancy] and by birth reaches significant levels, especially with respect to speech… If the newborn sensitivities to speech subserve bonding, discrimination of and preference for the maternal voice should be evident near birth.”

To test this hypothesis, ten newborn babies were tested for suckling activity, categorized in this study as “interburst interval” or IBI, using a pressure-sensitive transducer and nonnutritive nipple to activate either their mother’s or another mother’s voice. In eight out of ten infants, the

---

[75] “Broca area,” *Britannica.com* (online encyclopedia accessed October 14, 2023, https://www.britannica.com/science/Broca-area).

results showed that IBIs shifted by 34% in response to the mother's voice.[76] The researchers concluded that "the early preference demonstrated here is possible because newborns have auditory competencies adequate for discriminating individual speakers: They are sensitive to rhythmicity, intonation, frequency variation, and phonetic components of speech."[77]

These studies demonstrate the overall purpose of preferential discrimination. For a baby, external sounds are signals of safety and danger, as well motivators for action. Sounds can also be used to express emotions (crying, for example), to learn about one's surroundings, and to procure nourishment or connection. The involuntary gravitation towards something positive and familiar, such as a maternal voice, is the first instance of sound preference in an individual's life as well as the start of complex aural preferences. Gratification leads to expectation, and expectation leads to prediction; a cycle forms from repetitious behaviors seeking contentment. This perspective is supported by neurologists such as Carlota Pagès Portabella, who says that "when we listen to music, each sound we hear helps us to imagine what is coming next. It what we expect is fulfilled, we feel satisfied. But if not, we may be pleasantly surprised or upset."[78]

Survey: Introduction, Background, and Methods

In this section I describe a survey conducted during the Fall of 2023 among students at the University of Wisconsin-Madison titled, "Seeking a Range of Acceptability: The Effects of

---

[76] Anthony J. DeCasper and William P. Fifer, "Of Human Bonding: Newborns Prefer Their Mothers' Voices," *Science* 208.4448 (1980): 1175.

[77] DeCasper et al., "Of Human Bonding," 1176.

[78] "Musical Perception: Nature or Nurture?," *Neuroscience News* (accessed October 10, 2019).

Timbre, Volume, and Pitch on Musical Sound Preferences.” All materials pertaining to the study, including samples, questionnaires, diagrams, and photos, can be found either in the Appendix or online. The survey was approved by the Institutional Review Board.[79]

Sound preferences are different for everyone; as indicated above, people come to expect and favor certain sounds due to their upbringing, demography, cultural exposure, and so forth. This is reflected in the sheer diversity of the ways people express their tastes in the modern world, such as through musical genres, instruments, and individual compositions. Despite these variations, large-scale preferential patterns and trends can be observed in music that seem to indicate more consistency across the population than one might expect. The physiology of the human ear and the auditory cortex gives us innate sensitivity to qualities in sound such as pitch (frequency), volume (amplitude), and timbre (harmonic spectra). The way the auditory system connects to the memory and the limbic system result in many commonly held preferences and aversions.

A study by Jeroen Delplanque et al. studied the effects of complexity on aesthetic music preferences, using a variable series of seven grand piano tones (based on conventional Western tonality) to represent low and high complexity. The results of the study showed that participants generally favored series of sounds with intermediate complexity over those with too high or low complexity.[80] In a similar vein, a study by Trevor Cox sought to investigate the “horribleness” of sounds by surveying people through an interactive internet webpage. Thirty-four sounds were

---

[79] A compensation gift equivalent to $5 was offered to all participants, which is a factor that could potentially influence the resulting data pool. The filter questions (explained later) served to separate data from the participants who answered dishonestly (for the sake of receiving the gift quickly) from those who answered honestly.

[80] Jeroen Delplanque, Esther De Kiifm Clio Janssens, and Tom Verguts,”The Sound of Beauty: How Complexity Determines Aesthetic Preference,” *Acta Psychologica* 192 (2019): 146.

presented, from low and almost eerie noise to vomiting and a dentist's drill. His data showed that vomiting was the most repulsive sound on the list, followed by microphone feedback, scrapes, and high-pitched squeaks. [81] He comments that disgust is influenced by both the characteristics of the sound itself as well as certain social elements, such as whether the sound is associated with "an immoral or unfair act."[82] His results suggest that consistently high sharpness and roughness levels in sounds (referring to McDermott's definitions) in addition to unpredictability in volume, pitch, or timbre affects a sound's undesirability. Association with a negative event (such as vomiting or a crying baby) was also shown to be a factor influencing negative reception.

My initial hypothesis for the present study is that the more stable and harmonically simple sounds would be rated higher than sounds that are more unstable and spectrally dense. The basis for this prediction is that most stable and relatively simple sounds people hear throughout their lives (such as the voice of a reassuring mother, or tones from a musical instrument) are paired with positive feelings of safety or nourishment. While Delplanque's study suggests that intermediate complexity is aesthetically preferred over simplicity, he focuses on the perception of multiple tones of varying frequencies (which involves syntax) as opposed to a single tone.

My hypothesis was tested by asking students to rate similar sounds according to personal preference. Participants' rankings of the unfamiliar sounds in the survey should show a gradient of preferences reacting to the changing sonic characteristics of sounds throughout the questionnaire. During this survey, I collected responses anonymously from 147 students with UW-Madison email addresses. I placed two filter questions (which stated: "This is a filter

---

[81] Trevor Cox, "Scraping Sounds and Disgusting Noises," *Applied Acoustics* 69 (2008): 1119.
[82] Cox, Scraping Sounds and Disgusting Noises," 1119.

question: Please choose "X" for each of the four questions below and proceed") into the questionnaire to reduce overall survey "noise" and minimize the impact of dishonest submissions on the data set; data containing incorrect answers to either question would be removed from consideration. Out of 147 total participants, 71 passed both filter questions. The students were asked to wear headphones for the duration of the survey but were free to use any electronic device with a sound system. The survey presented a total of 16 groups of sounds, each group containing four individual tones. Within each group, students were asked to listen to each of the four samples and rate them according to their preference on a scale from 1 to 5, replaying as often as they wished. The precise prompt for each question was as follows: "How would you rate your enjoyment of the first sound? 5 = Very Much | 3 = Neutral | 1 = Very Little."

My survey focused particularly on the influence that timbre, pitch, and volume have on perceptions of sound. To this end, I generated sounds in which pitch and volume ranged across a spectrum from "stable" to "unstable"; timbre was altered to vary progressively in harmonic richness. The sound samples were created using an analog MiniMoog Voyager Synthesizer, and I have included photos of the exact settings in the appendix. Each of the qualities of timbre (T), pitch (P), and volume (V) varied across four levels of intensity. The details for each individual quality and intensity level can be found in the tables below. I chose these parameters because they can be easily and somewhat linearly modified on a Moog synthesizer or audio software, and because each quality can potentially hinder enjoyment when taken to either a high or low extreme. I used vibrato to represent stability/instability in P because vibrato is a commonly used technique in music to alter the pleasantness of a sustained tone. I varied V according to my informal judgement of what might be considered stable or unstable to most people. I chose to alter the partials present in T by using a filter dial on the Moog synthesizer because this type of timbral augmentation is widely used in concerts by performing artists today. The purpose of using these parameters was to create a range of similar sounds that, when presented to the public, could produce a range of responses that can fluctuate according to personal preferences. Careful consideration was given to the intensity ranges of each quality to ensure that no sound was obviously and universally repulsive.

Timbre: Timbre varied according to partial number; a filter dial on the synthesizer was used to manipulate how many partials were present in the sound wave. The initial sound most closely resembles a sawtooth wave, and one oscillator was used. Rather than using precise computational methods during the filtering process, I used my ears to judge which partials were present in the final samples. Appendix 2 displays the spectrograms of each level of T.

Level 1	Level 2	Level 3	Level 4
Fundamental + First partial	Previous level + 2nd and 3rd partials	Previous level + 4th and 5th partials	Previous level + 6th and 7th partials

Pitch: Pitch was varied by vibrato distance, which was approximately set to 4.5 oscillations per second. The maximum vibrato distance ranged from a semitone below to a semitone above the original pitch. An analog modular wheel was used to vary the pitch fluctuation levels of the vibrato.

Level 1	Level 2	Level 3	Level 4
Modular wheel at 0%	Modular wheel at ~33%	Modular wheel at ~66%	Modular wheel at 100%

Volume: The volume envelope for each sound was adjusted during post-production. I used personal discretion while creating the envelope function, which are shown in the appendix; percentages listed in this table are approximate.

Level 1	Level 2	Level 3	Level 4
Volume unaltered from the original sample	Volume ranges from ~75% to ~125% of the original	Volume ranges from ~50% to ~150% of the original	Volume ranges from ~25% to ~175% of the original

Individual sounds were named according to their respective qualities and intensities. For example, T1P3V4 means that the timbre was set to level 1, the pitch to level 3, and the volume to level 4. A total of 64 individual sounds were created this way, to account for every possible variation, with each sample spanning 3 seconds. All sounds were randomly placed into groups of 4 and set to play consecutively, with 2 seconds of silence in between, resulting in a total of about 20 seconds per group. The groups were randomly arranged as follows:

Group Name	**Sample 1**	**Sample 2**	**Sample 3**	**Sample 4**
Group 1	T2 P2 V4	T4 P4 V2	T3 P3 V2	T1 P2 V3
Group 2	T2 P3 V1	T1 P3 V4	T3 P4 V3	T2 P1 V4
Group 3	T4 P2 V3	T3 P1 V1	T4 P1 V2	T2 P4 V4
Group 4	T1 P4 V1	T4 P3 V4	T1 P1 V3	T3 P2 V1
Group 5	T3 P3 V3	T2 P2 V1	T1 P2 V2	T3 P1 V4
Group 6	T1 P3 V2	T4 P2 V1	T2 P4 V2	T4 P3 V1
Group 7	T3 P4 V1	T4 P4 V3	T2 P1 V2	T1 P4 V4
Group 8	T2 P3 V4	T3 P2 V3	T1 P1 V1	T4 P1 V3
Group 9	T4 P3 V2	T1 P2 V1	T3 P4 V4	T2 P2 V3
Group 10	T1 P3 V3	T2 P1 V1	T3 P3 V4	T4 P2 V4
Group 11	T3 P2 V2	T2 P3 V3	T1 P1 V4	T4 P4 V4
Group 12	T4 P1 V1	T2 P4 V3	T3 P4 V3	T1 P4 V2
Group 13	T4 P2 V2	T3 P1 V2	T2 P2 V2	T4 P3 V3
Group 14	T1 P3 V1	T2 P3 V4	T3 P1 V3	T1 P2 V4
Group 15	T2 P1 V3	T4 P4 V1	T4 P1 V4	T3 P2 V4
Group 16	T1 P1 V2	T3 P3 V1	T2 P4 V1	T1 P4 V3

Sounds are available online, with links in the footnotes.[83]

---

[83] Sounds are available on SoundCloud, following this link: https://soundcloud.com/willey-lee?utm_source=clipboard&utm_medium=text&utm_campaign=social_sharing

Results and Discussion

Most Preferred Sounds by Average

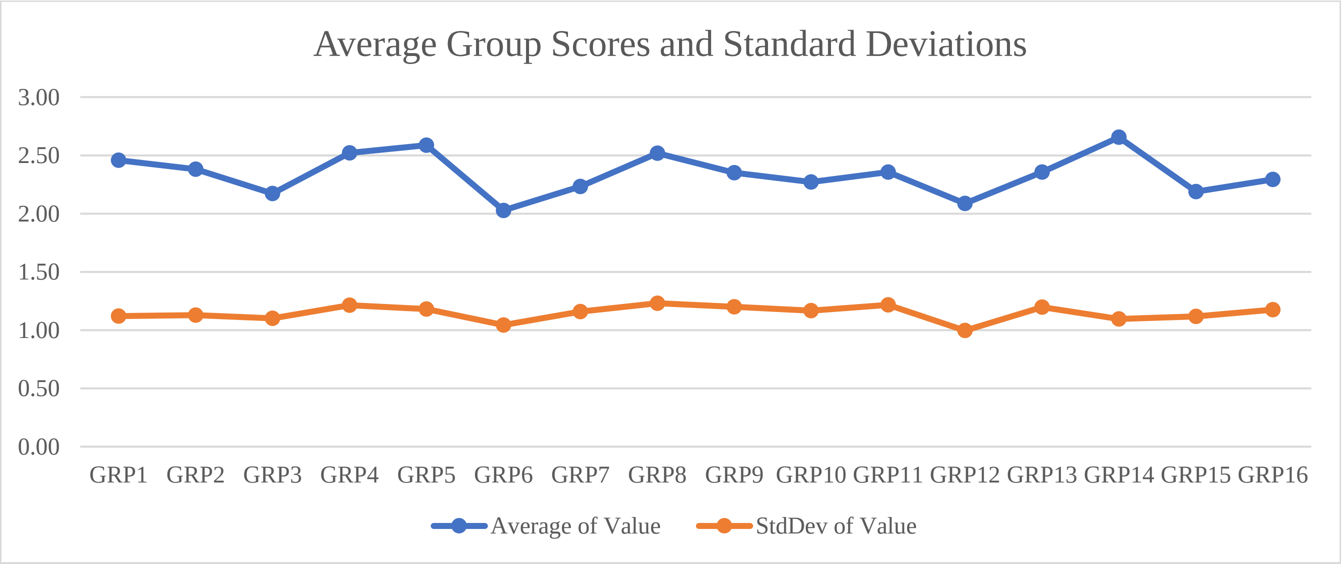
Group Number	Sound Sample	Average
Group 16	T1P1V2	3.51
Group 4	T1P1V3	3.51
Group 8	T1P1V1	3.48

Least Preferred Sounds by Average

Group Number	Sound Sample	Average
Group 11	T4P4V4	1.45
Group 6	T4P3V1	1.48
Group 7	T4P4V3	1.51

Average Scores and Standard Deviations Across All Groups

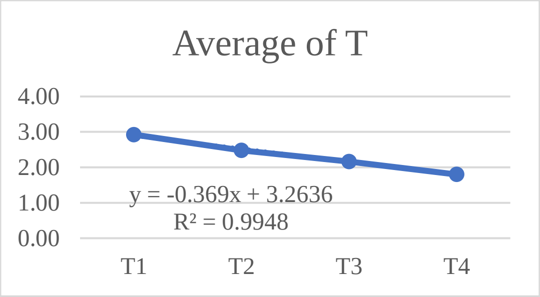
Group Labels	Average Score	StdDev of Score
Group 1	2.46	1.12
Group 2	2.38	1.13
Group 3	2.17	1.10
Group 4	2.52	1.21
Group 5	2.59	1.18
Group 6	2.03	1.04
Group 7	2.23	1.16
Group 8	2.52	1.23
Group 9	2.35	1.20
Group 10	2.27	1.17
Group 11	2.36	1.22
Group 12	2.09	1.00
Group 13	2.36	1.20
Group 14	2.65	1.10
Group 15	2.19	1.12
Group 16	2.29	1.18



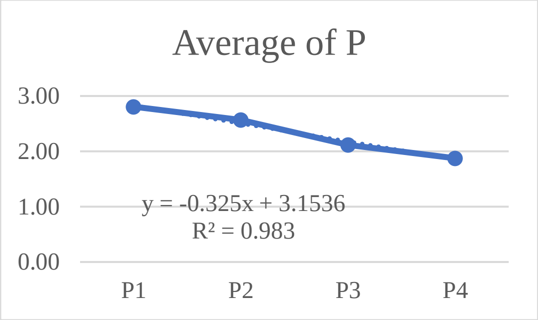
Group averages were moderately consistent and hovered between 2.03 and 2.65, with dips in Groups 3, 6, and 12. This could be due to the relatively high concentrations of levels 3 and 4 in T and P (trends seen in T and P will be discussed later). Standard deviations generally stayed steady between 1.00 and 1.23. The standard deviations support the idea that the survey produced consistent results, and they also appear to show that there was little attention deterioration as the survey progressed.

Average scores of T, P, and V

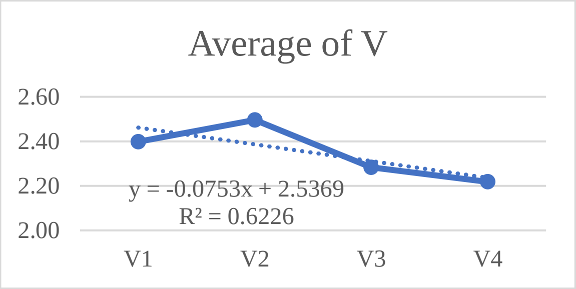
Timbre Level	Average Score
T1	2.93
T2	2.48
T3	2.16
T4	1.8



Pitch Level	Average Score
P1	2.81
P2	2.57
P3	2.11
P4	1.88



Volume Level	Average Score
V1	2.4
V2	2.5
V3	2.28
V4	2.22



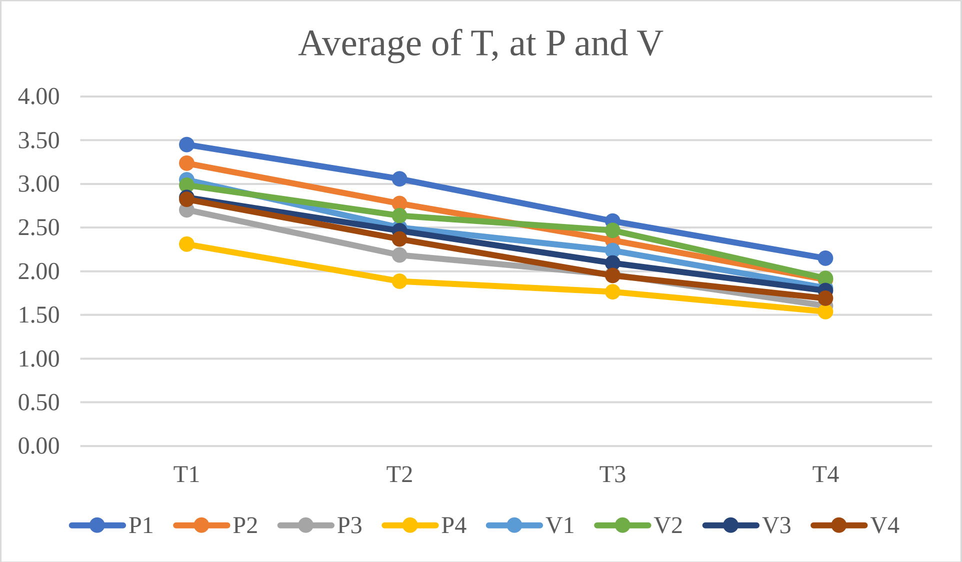
Average scores in T and P decreased as the level of intensity increased. The coefficient of determination, R^2, for the averages of T and P were 0.99 and 0.98, showing a strong negative correlation between T/P and average scores. These figures support my initial expectation that participants would favor sounds that are more stable or harmonically simple. Additionally, the rate of decrease in average scores as T or P increased was consistent, as shown by the R^2, and indicates a progressive preference with regards to intensity levels (i.e., T2 was more popular than T3, and T3 was more popular than T4). The strength of the correlation is salient because of the nature of the survey; samples were randomized with regards to presentation order and exhibited qualities, making every ranking unique to each participant's preferences.

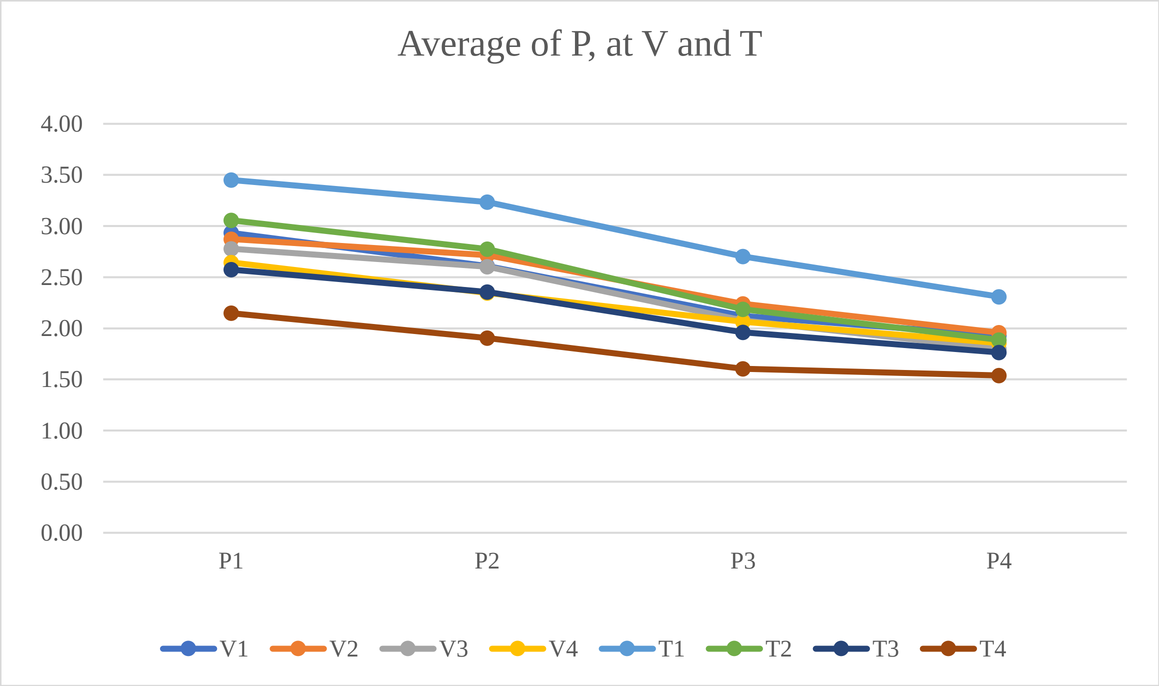
One exception was V, which did not show as strong of a correlation; R^2 for V was 0.62, and V2 was more popular than V1. This result is surprising, given the original hypothesis, but it aligns with common experiences. While stable sounds can be pleasant or calming, they can also be boring or uninspiring; this data seems to support Carl Seashore's claim, as I mentioned in a previous section, that humans are accustomed to periodic pulsations in sound.[84] However, this perspective might seem to imply that P2 (containing only a small amount of pitch vibrato) to perform better than P1, which was not the case. One possible factor is cumulative "annoyance," or disdain built up over time; after hearing a string of samples with varying levels of vibrato, P1 might have served as a moment of respite. Another possibility is that the increase in intensity from P1 to P2 is more perceptually salient than the equivalent in V. I selected the sonic parameters of T, P, and V at my own discretion, so it is possible that the perceptual difference between P1 and P2 is greater than the difference between V1 and V2. Alternatively, it may be that subtle variations in volume are in general less likely to be noticed consciously and create

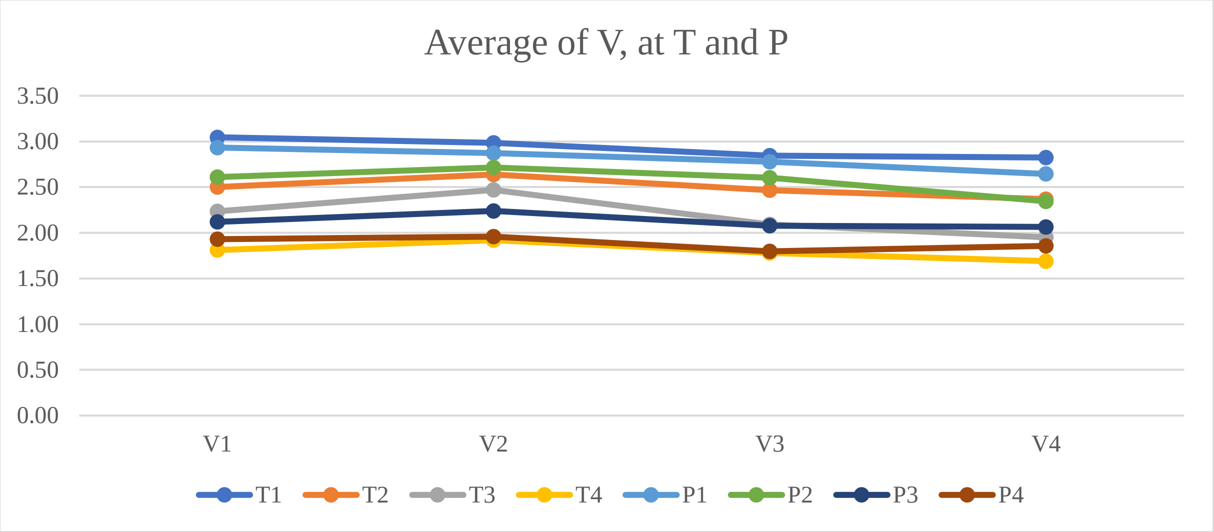
---

[84] Seashore. "The Psychology of Music," 31.

only a subliminal feeling that a sound is more “natural.” These possibilities are further magnified by the survey’s uncontrolled environment; since environmental variables were not monitored, and the headphones themselves were not administered by the study team, stimuli from a participant’s physical and sound setting (i.e. the sounds that could have been heard through headphones) could have impacted a participant’s choice patterns.



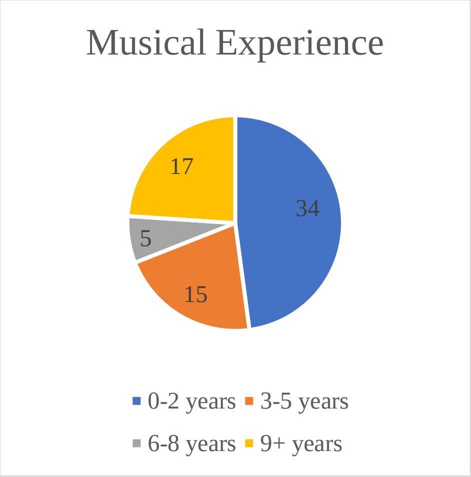




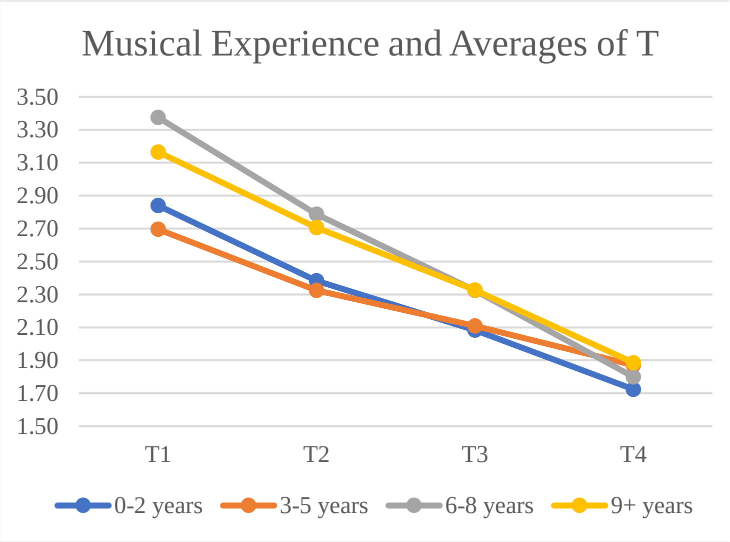
The graphs above further visualize the negative correlations observed in T and P. The graph of V shows that changes in V scarcely impacted the scores of T or P.

Musical Experience and Averages

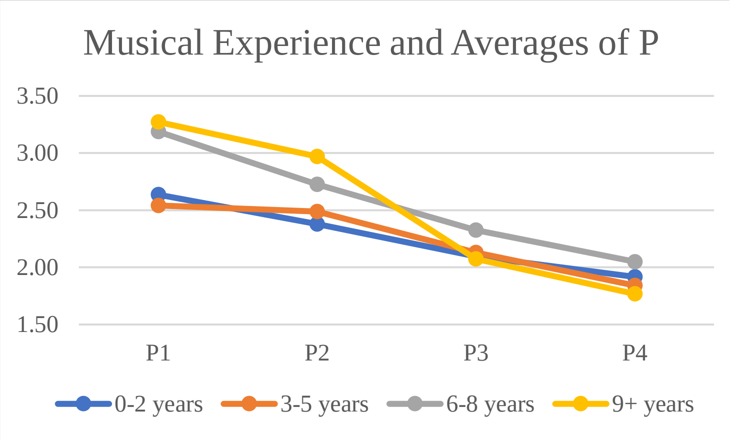
Musical Experience	# of Responses
0-2 years	34
3-5 years	15
6-8 years	5
9+ years	17



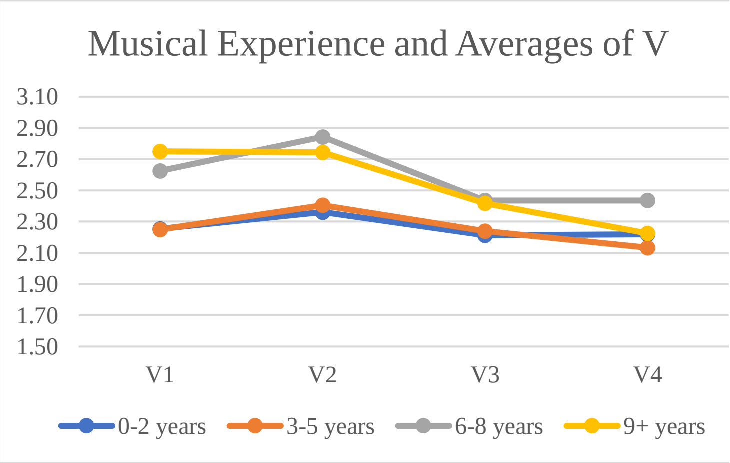
Musical Exp.	**T1**	**T2**	**T3**	**T4**
0-2 years	2.84	2.38	2.08	1.72
3-5 years	2.70	2.33	2.11	1.87
6-8 years	3.38	2.79	2.33	1.80
9+ years	3.17	2.71	2.33	1.89



Musical Exp.	**P1**	**P2**	**P3**	**P4**
0-2 years	2.64	2.38	2.10	1.92
3-5 years	2.54	2.49	2.13	1.84
6-8 years	3.19	2.73	2.33	2.05
9+ years	3.27	2.97	2.07	1.77



Musical Exp.	**V1**	**V2**	**V3**	**V4**
0-2 years	2.25	2.36	2.21	2.22
3-5 years	2.25	2.40	2.24	2.13
6-8 years	2.63	2.84	2.44	2.44
9+ years	2.75	2.74	2.42	2.22



The data show that musical experience factors into perceptions of T and P. Across T, P, and V, participants with more than six years of musical experience selected noticeably higher scores for levels one and two. Those with at least nine years of experience indicated noticeably high levels of dislike for P3 when compared to P2, which may correspond with specialized and developed tastes in addition to more frequent exposure to music over time. While this does appear to support the notion that musical experience contributes positively towards the enjoyment of sounds, it is important to note that these two groups only comprise 31% of the data set (22 students out of 71). A larger sample size with each group sufficiently represented would better support the validity of these observations.

Conclusion: The "Soulful Sound" Framework

Based on the research described above, I propose a loose framework for understanding human sound preferences. I use the term "soulful sound" to convey the way in which people prefer sounds that are familiar, relatively easy to remember, contain recognizable elements of pitch and timbre, and are contextually appropriate. Some of these traits can also describe unliked sounds to a certain extent, such as those used in Cox's study, but I suggest that the difference lies in the emotional association created due to external reasons such as social etiquette or connection to physiological or psychological discomfort (for example, an ambulance or police siren). Along these lines, the survey data appear to suggest that, at least in the domains of timbre and pitch, simplicity and regularity are overall more significant to the positive reception of sounds than great complexity. This result makes sense, given what I discussed about noise; an unpredictable sound with greater spectral density might be displeasing because the mind cannot easily process the sound or associate the sound with a positive emotion. It could also be that most sounds that warn us of danger tend to exhibit spectral complexity and contain erratic elements.

The chronometric model proposed by Elvira Brattico et al. provides a potential outline for the order of neural processes during the formation of preferences. Figure 3 details the sequence from sound to taste: Early emotional responses starting from the brainstem begin the process of perception. These automatic reflexes, such as the startle reflex (triggered by surprising sounds), influence initial arousal and "subliminally may determine actions and judgments, as demonstrated by priming paradigms."[85] Stylistic and syntactical judgements follow early emotional responses, eventually leading into aesthetic conclusions as well as the development of a broad musical taste. Taste is defined as "a long-term set of preferences, aesthetic judgments, values, and attitudes," and is formed by numerous instances of "conscious liking," with which there is a feedback loop (i.e., taste informs conscious liking, and conscious liking begets taste).[86]

A sound is "soulful" when it or a similar sound has been associated with positive emotions; people enjoy sounds resembling ones they enjoyed before. I borrowed the usage of the word "soul" from Forkel, who commented that soul was a desirable quality more present in the clavichord's sound than in the harpsichord's. J.S. Bach might have favored the mellow and relatively simple sound of the clavichord over the bright and spectrally complex sound of the harpsichord (Figures 4 and 5 show a rough spectral comparison), or the clavichord's sound profile may have more-closely resembled those of other popular instruments such as the lute or violin. He also might have been referring to the clavichord's ability to bend pitches (using a technique known as *Bebung*) and exercise greater control over dynamics, features that the human voice possesses and the harpsichord lacks (especially considering the importance of familiarity in

---

[85] Elvira Brattico, Brigitte Bogert, and Thomas Jacobsen, "Toward a Neural Chronometry for the Aesthetic Experience of Music," *Frontiers in Psychology* 4 (2013): 6.
[86] Brattico et al., "Towards a Neural Chronometry," 13.

forming tastes). Concordantly, my survey suggests that a small amount of variation in volume might be generally desirable in a musical tone.

The "feedback loop" mentioned above leads to a complex preferential web that creates expectations for future stimuli. When an expectation's borders are exceeded (sonically or aesthetically), the limbic system can potentially activate "sympathetic activity" that reflects "the intensity of psychological (emotional) activation."[87] This can lead to terror or fright, Bjorn Merker notes, but also to "a tamed version of the same in the form of being 'touched,' 'moved,' 'impressed,' and… even 'awed' by what is heard."[88]

For a musician, one primary goal is to perform music in a way that leads the listener into a state of awe. Performers in the discipline of classical music spend countless hours perfecting their sounds to make them more appealing, emotionally impactful, and ultimately more soulful to the listener. However, the issue of "why" a sound is considered beautiful is a difficult subject that is rarely discussed with precision apart from a cultural context. As I mentioned, the positive reception of a sound is the result of an ongoing cycle of experiences and memories that are strengthened through repetition; this repetition often leads to the generation of feelings of nostalgia and the evocation of strong emotions. According to Brattico, "elicitation of memories and nostalgia are listed among the main reasons for listening to music and for the strongest bodily changes in both elderly and young adults."[89] While speech, in particular, contains specific information that allows for relatively unambiguous and accurate communication, musical sounds do not have the same concrete meanings attached to them; one's memories, culture, personality,

---

[87] Bjorn Merker, "When Extravagance Impresses," *The Oxford Handbook of Music and the Brain*, ed. Michael Thaut and Donald A. Hodges (Oxford: Oxford University Press, 2019), 72.

[88] Ibid., 73.

[89] Brattico, "Towards a Neural Chronometry," 11.

and education, among many other factors, heavily influence the perception of sounds. Whether simple or complex, all preferences are derived from the biological systems that promote the goals of vitality and safety, and from our intellectual desire for stimuli that balance comprehensibility, predictability, and familiarity against novelty and intricacy.

Figure 3: From Elviro Brattico et al., "Toward a Neural Chronometry for the Aesthetic Experience of Music," *Frontier in Psychology* 4.206 (2013), 2.

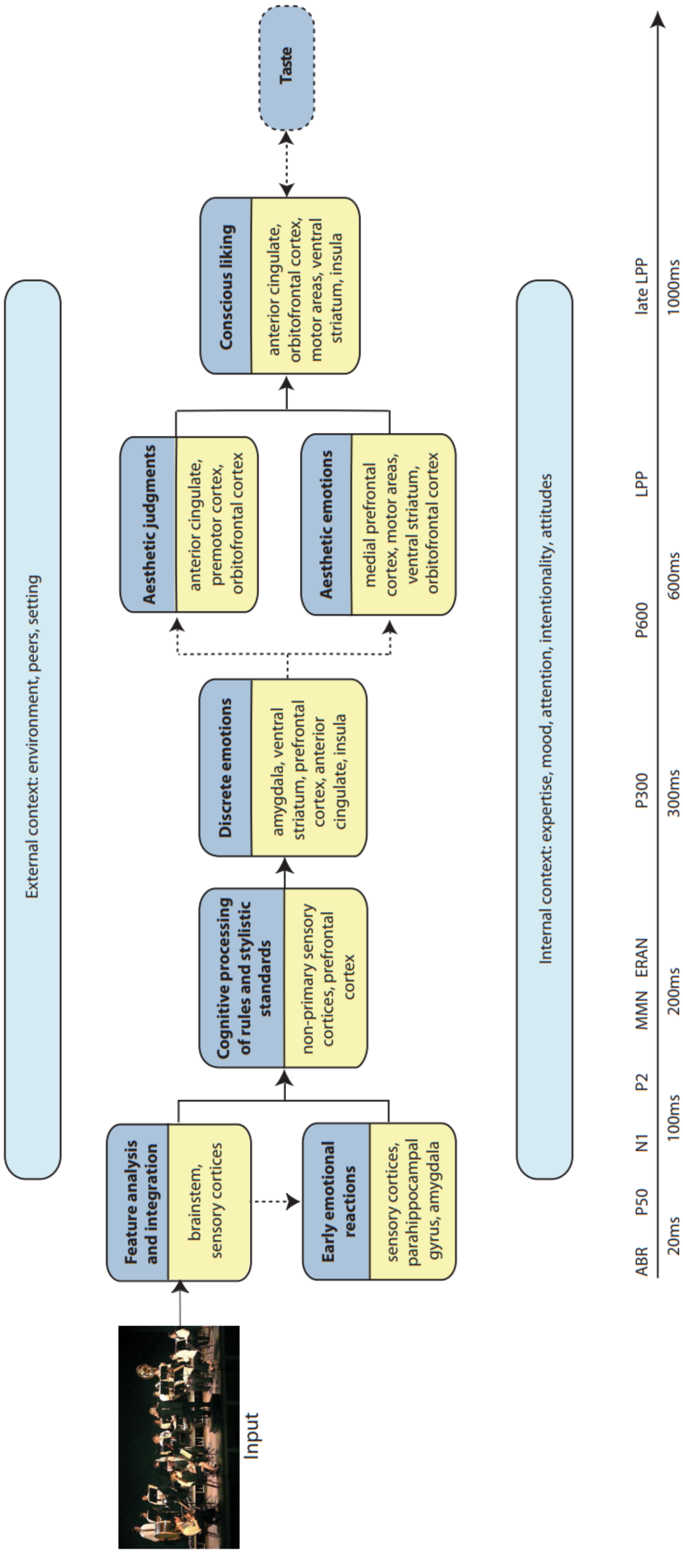


Figure 4: Spectrogram of a Harpsichord Playing D3

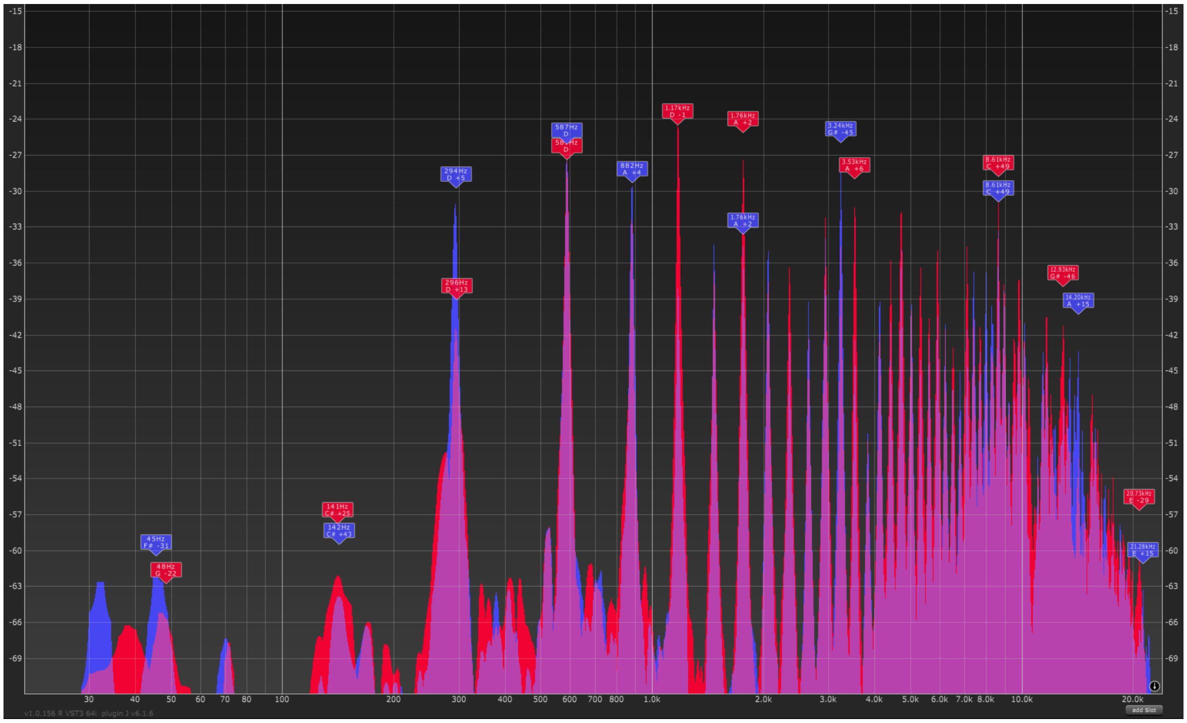
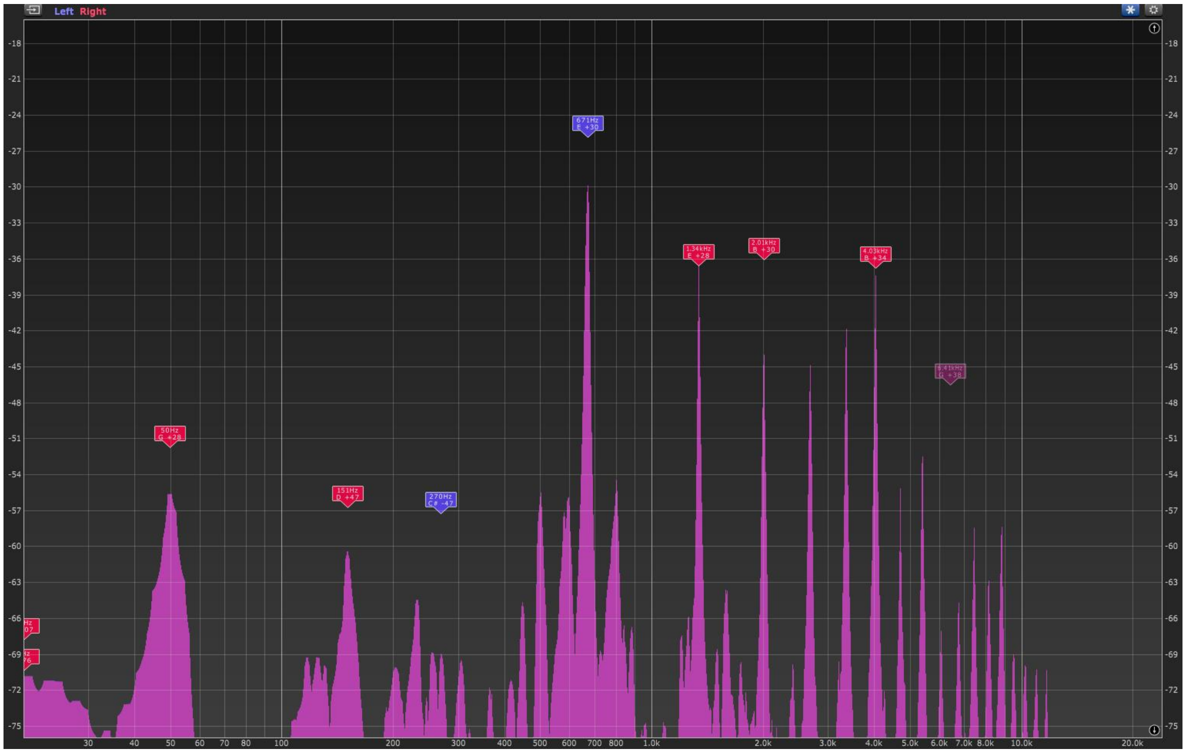
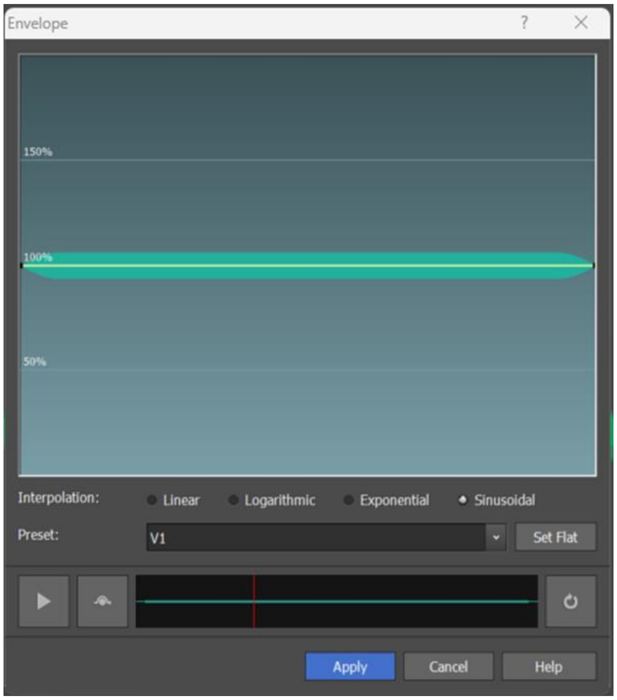


Figure 5: Spectrogram of a Clavichord Playing E5



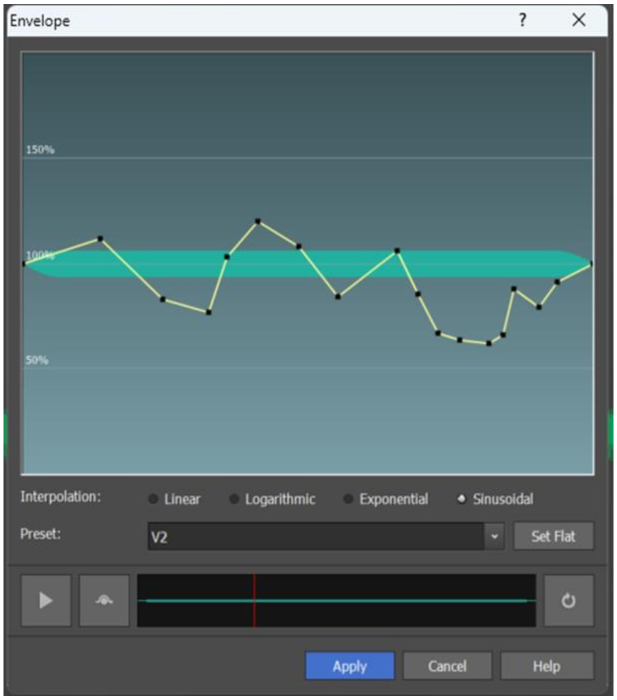
Appendix 1: Volume Envelopes

The following diagrams are screenshots taken of the volume envelope templates applied to all samples. WavePad Audio Editor was used to master these synthesizer samples post-production. Presets for V1, V2, V3, and V4, shown below, were overlaid on top of each original audio clip. The lines and nodes indicate general volume flow and shape across the waveform, and sinusoidal interpolation was used to create natural volume shifts between nodes. All modifications were made to the base audio sample (V1), as opposed to progressively augmenting each new sample (for example, preset V3 was NOT applied to the already-augmented waveform of V2, which would cause a much more drastic change).



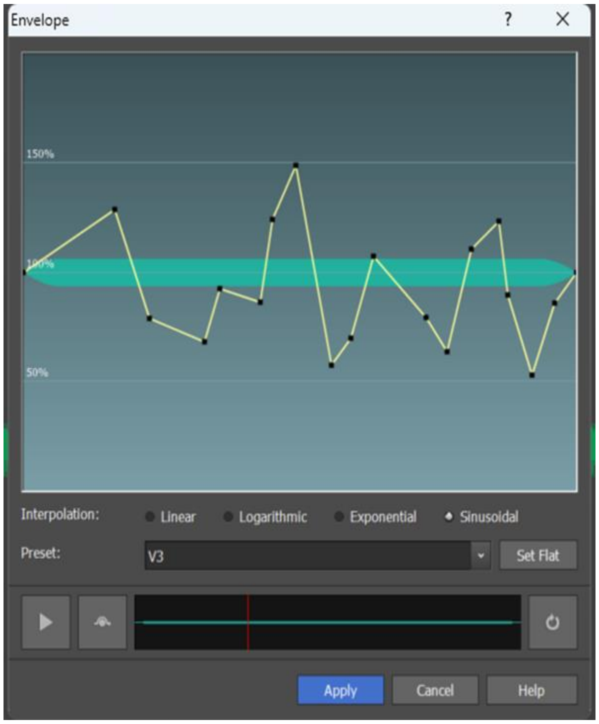
Interpolation: ○ Linear ○ Logarithmic ○ Exponential ● Sinusoidal

Preset: V1



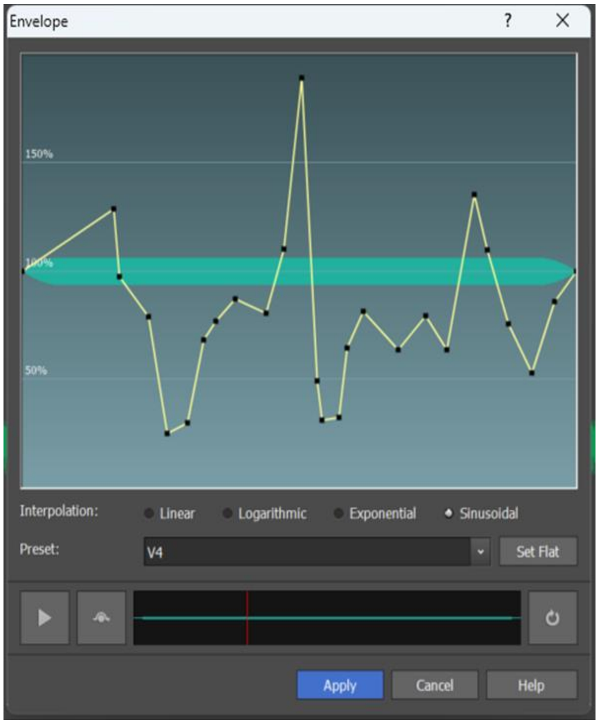
Interpolation: ○ Linear ○ Logarithmic ○ Exponential ● Sinusoidal

Preset: V2



Interpolation: ○ Linear ○ Logarithmic ○ Exponential ● Sinusoidal

Preset: V3



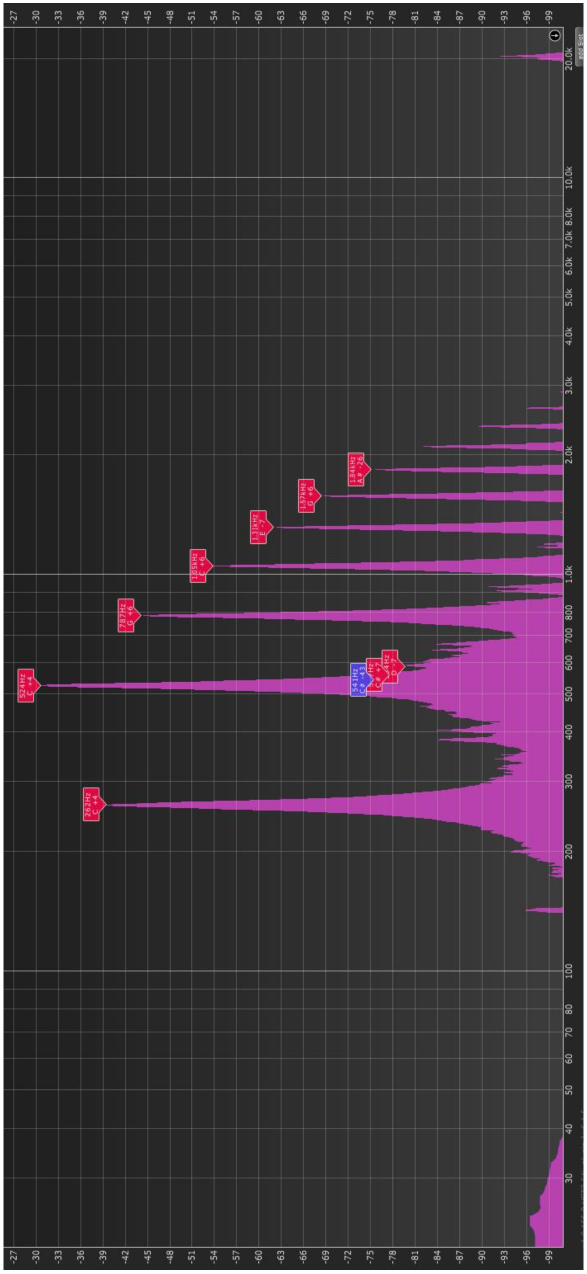
Interpolation: ○ Linear ○ Logarithmic ○ Exponential ● Sinusoidal

Preset: V4

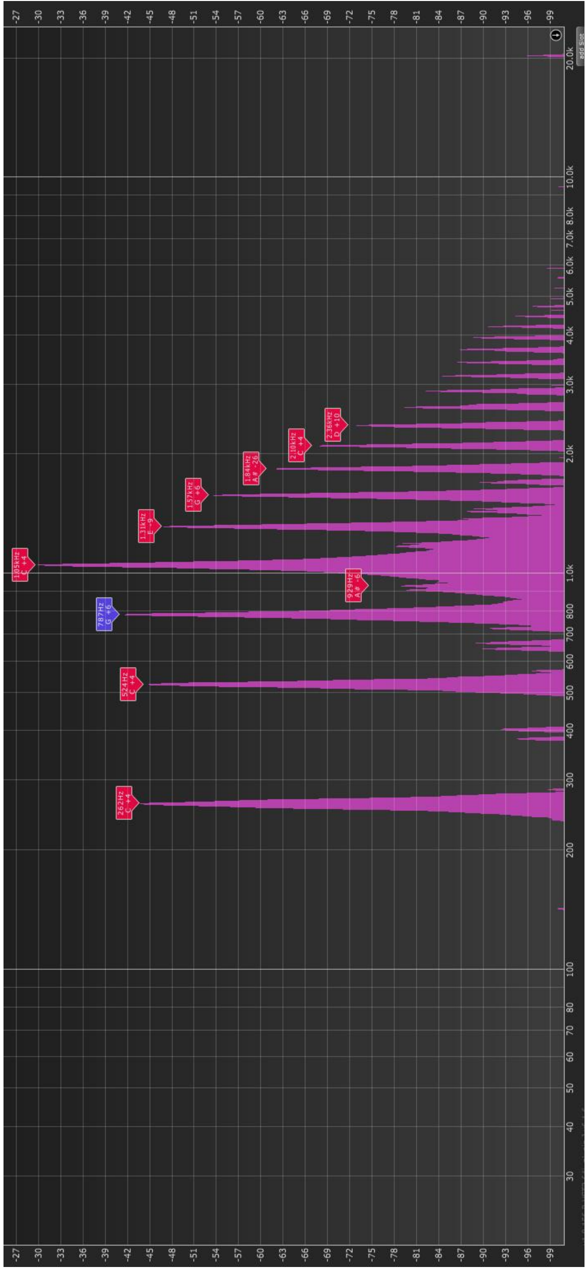
Appendix 2: Spectrograms of T and P

Spectrograms of T were generated using SpectrumAnalyzer by SIR Audio Tools. Spectrograms of P were generated using Spectrogram by sciencemusic.org (https://spectrogram.sciencemusic.org).

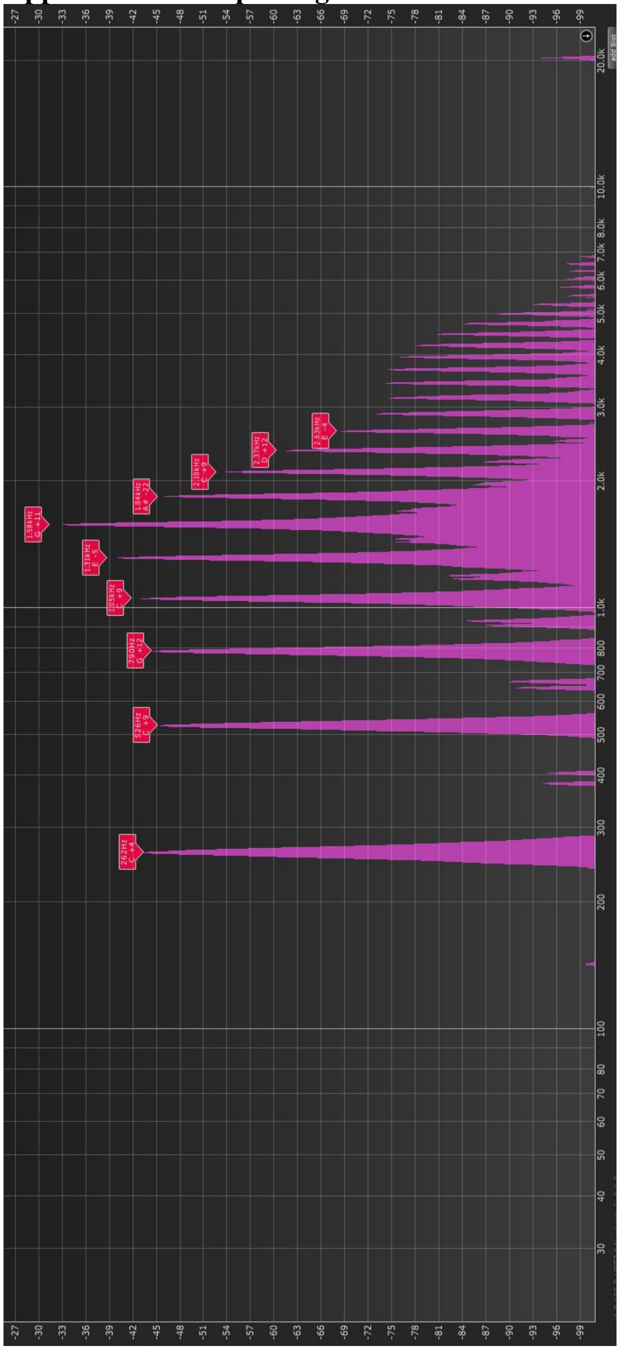
Appendix 2a: T1 Spectrogram



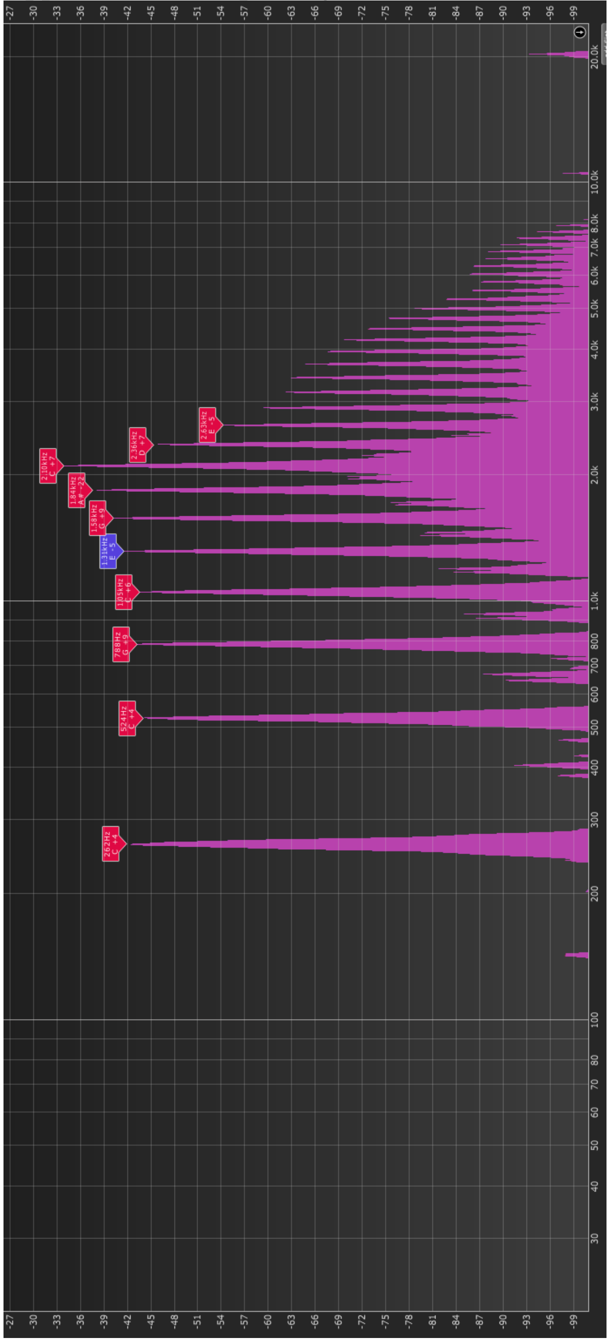
Appendix 2b: T2 Spectrogram



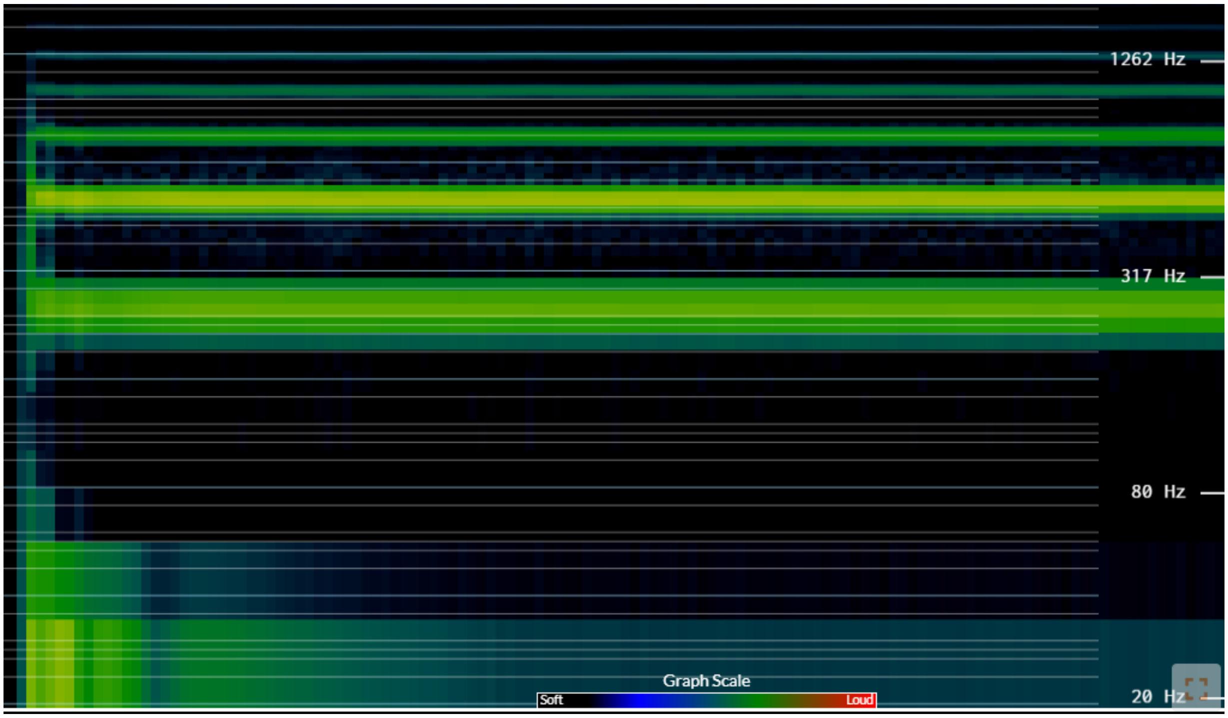
Appendix 2c: T3 Spectrogram



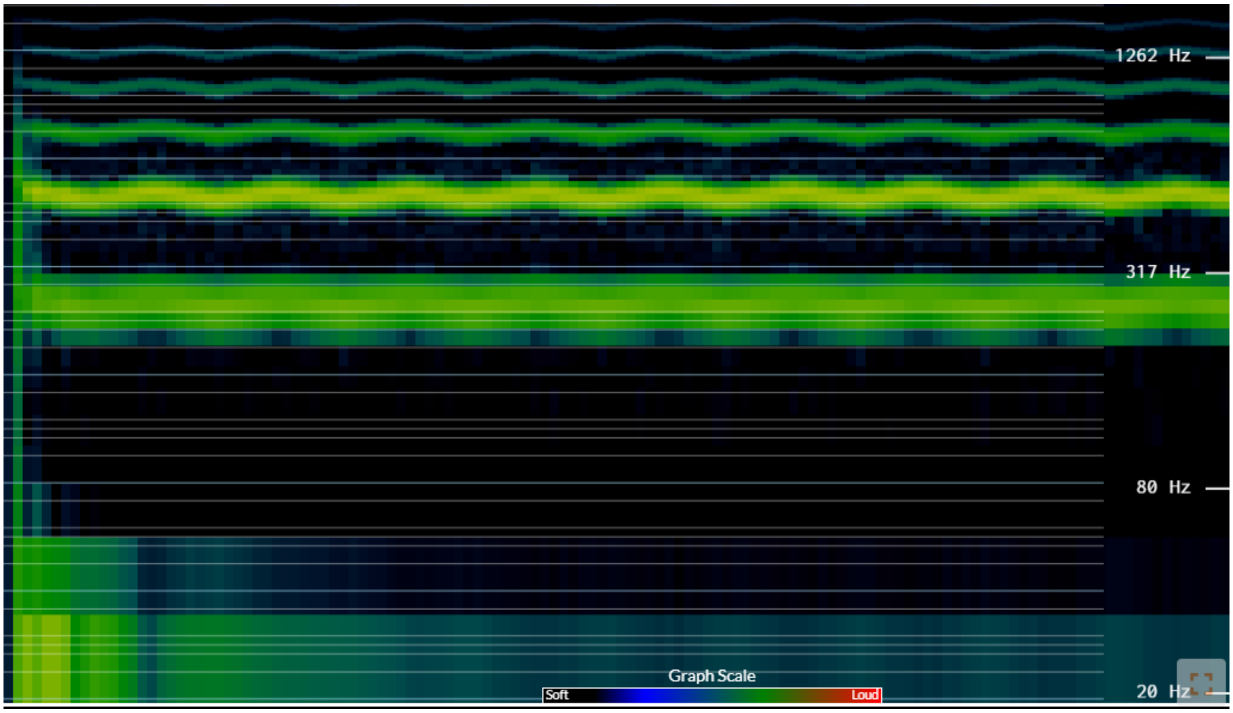
Appendix 2d: T4 Spectrogram



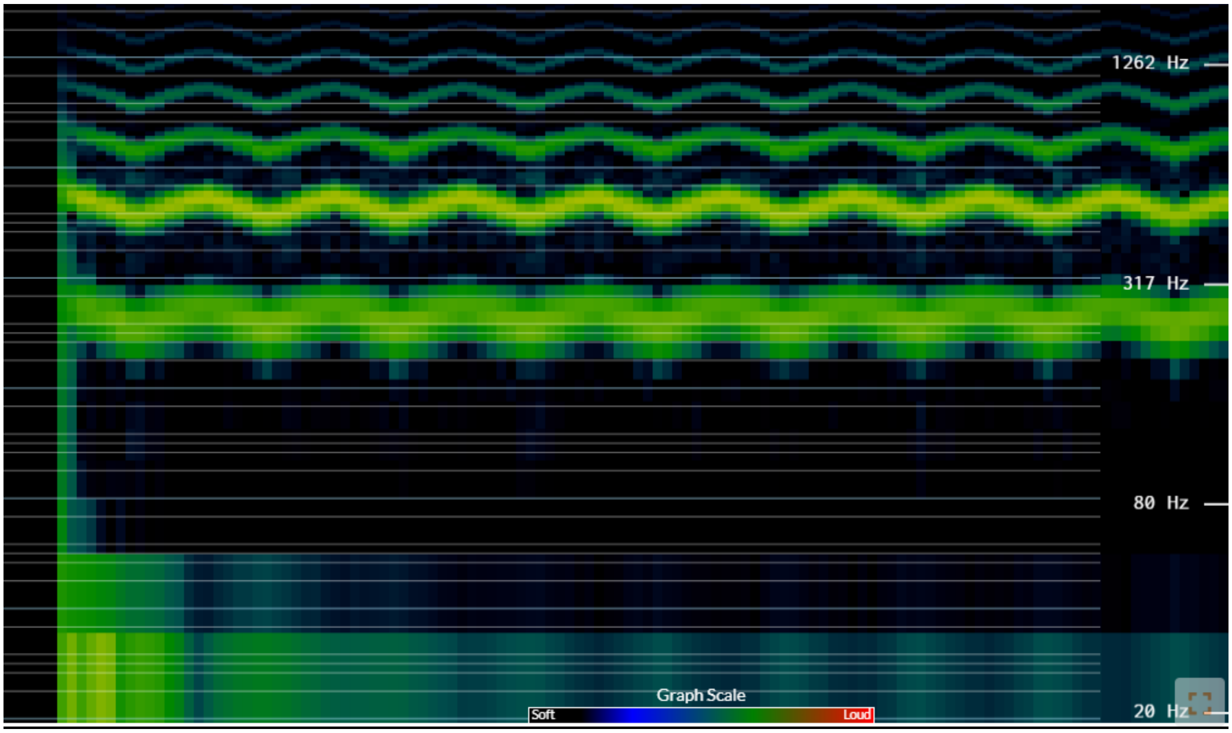
Appendix 2e: Spectrogram of P1 (with T1)



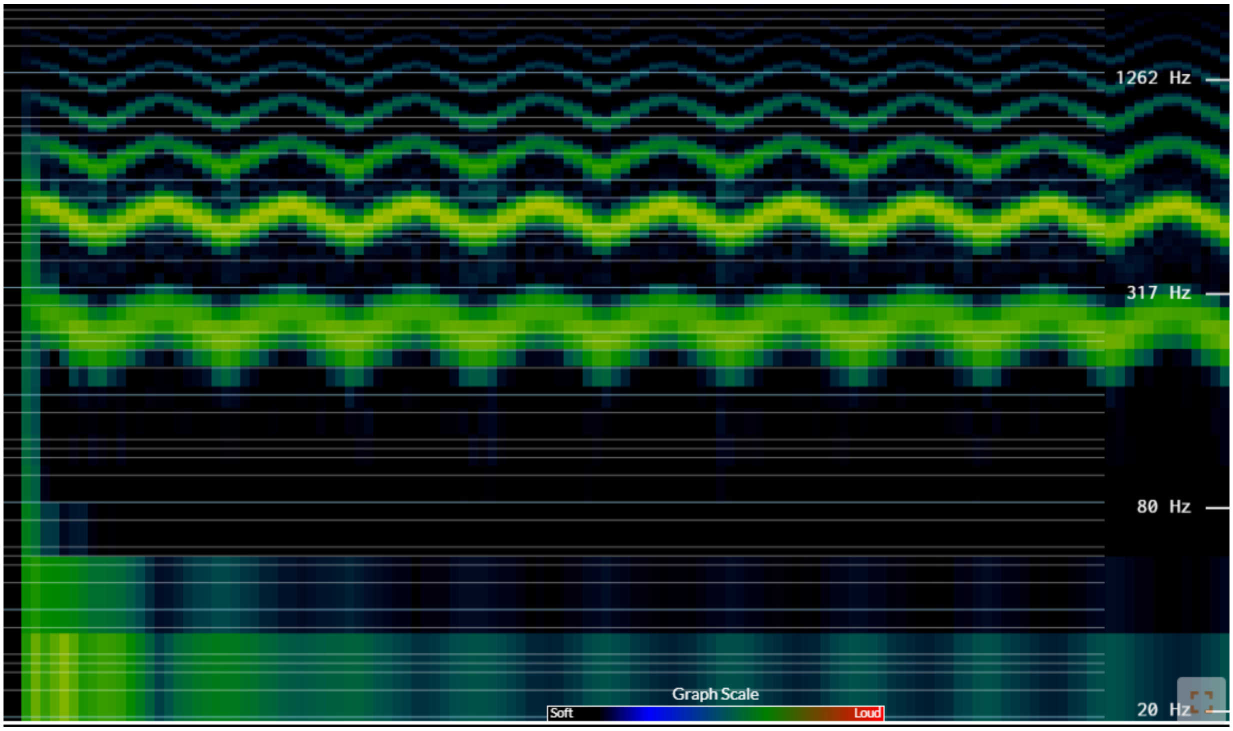
Appendix 2f: Spectrogram of P2



Appendix 2g: Spectrogram of P3



Appendix 2h: Spectrogram of P4



Appendix 3: Settings Used on the MiniMoog Voyager

Each picture is labeled according to the corresponding sound quality and level. Some additional is given, when applicable.

Appendix 3a: Setting Used for T1

Appendix 3b: Setting Used for T2

Appendix 3c: Setting Used for T3

Appendix 3d: Setting Used for T4

Appendix 3e: Setting Used for P1

Appendix 3f: Setting Used for P2

Appendix 3g: Setting Used for P3

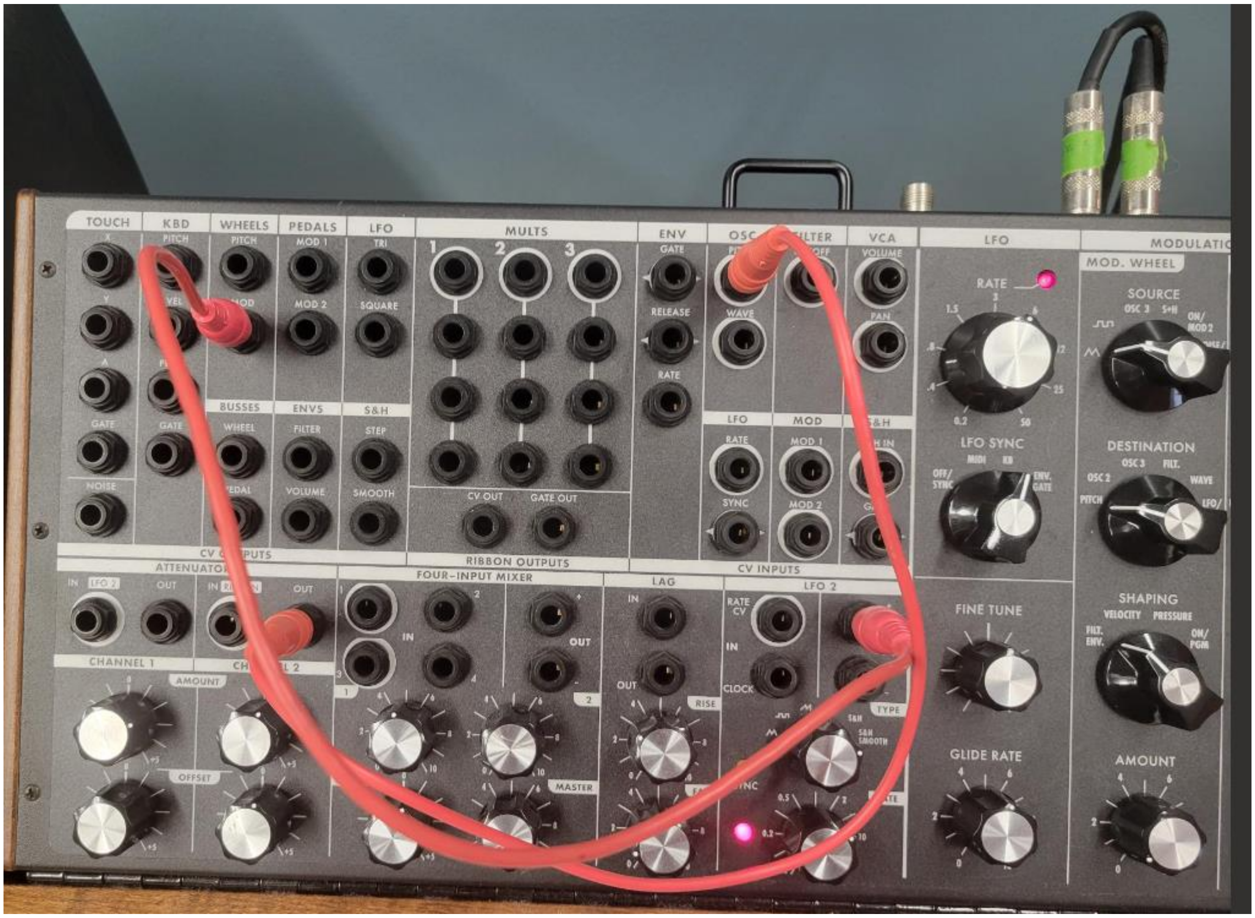
Appendix 3h: Setting Used for P4

Appendix 3i: Settings Used for Pitch Oscillation

(LFO rate influences vibrato oscillation per second)

Appendix 3j: Miscellaneous Settings


LFO
MODULATION BUSSES
MOD. WHEEL
PEDAL/ON
1
2
RATE
1.5
3
6
12
25
50
0.2
.4
LFO SYNC
MIDI
KB
OFF/ SYNC
ENV. GATE
FINE TUNE
SOURCE
OSC 3
S+H
ON/ MOD 2
NOISE/ PGM
DESTINATION
OSC 3
FILT.
OSC 2
WAVE
PITCH
LFO/ PGM
SHAPING
VELOCITY
PRESSURE
FILT. ENV.
ON/ PGM
OCTAVE
8'
4'
16'
2'
32'
1'
WAVE
4
6
2


TOUCH
KBD
WHEELS
PEDALS
LFO
MULTS
ENV
OSC
FILTER
VCA
LFO
MODULATIC
BUSSES
ENVS
S&H
CV OUTPUTS
ATTENUATORS
RIBBON OUTPUTS
FOUR-INPUT MIXER
LAG
CV INPUTS
LFO 2
CHANNEL 1
AMOUNT
OFFSET
RATE
LFO SYNC
FINE TUNE
GLIDE RATE
MOD. WHEEL
SOURCE
DESTINATION
SHAPING
AMOUNT


OUT
TYPE
S&H
S&H SMOOTH
RATE
0.5
2
10
20
FILT. ENV.
PRESSURE
ON/ PGM
GLIDE RATE
4
6
2
8
0
10
AMOUNT
4
6
2
0
10
VELOCITY
FILT. ENV.


NOISE
ON
OSCILLATOR 3
OSCILLATOR 2
OSCILLATOR 1
EXTERNAL
DUAL LOWPASS
HIGHPASS / LOWPASS
KB. CONT. AMOUNT
RESONANCE
SPACING
20
60
250
12K
4K

Bibliography

Brattico, Elvira, Brigitte Bogert, and Thomas Jacobsen. “Toward a Neural Chronometry for the Aesthetic Experience of Music.” *Frontiers in Psychology* 4 (2013): 6–13.

Britannica.com (online encyclopedia https://britannica.com).

Caclin, Anne, Elvira Brattico, Mari Tervaniemi, Risto Näätänen, Dominique Morlet, Marie-Hélène Giard, and Stephen McAdams. “Separate Neural Processing of Timbre Dimensions in Auditory Sensory Memory.” *Journal of Cognitive Neuroscience* 18.12 (2006): 1959–1972.

Cogan, Robert. *Sonic Design: The Nature of Sound and Music.* Cambridge: Publication Contact International, 1984.

Cogan, Robert. *New Images of Musical Sound*. Cambridge: Harvard University Press, 1984.

Campbell, Murray. “Partial.” Grove Music Online (accessed October 14, 2023).

Cox, Trevor. “Scraping Sounds and Disgusting Noises.” *Applied Acoustics* 69 (2008): 1195-1204.

Cowan, Nelson. “What Are the Differences Between Long-Term, Short-Term, and Working Memory?” *Progress in Brain Research* 169 (2008): 326.

Darwin, Charles. *Expression of the Emotions in Man and Animals* (1982). London: John Murray.

DeCasper, Anthony J., and William P. Fifer. “Of Human Bonding: Newborns Prefer Their Mothers’ Voices.” *Science* 208.4448 (1980): 1174–1176.

Delplanque, Jeroen, Esther De Loof, Clio Janssens, and Tom Verguts. “The Sound of Beauty: How Complexity Determines Aesthetic Preference.” *Acta Psychologica* 192 (2019): 146-152.

Eschrich, Susann, Thomas F Münte, and Eckart O Altenmüller. “Unforgettable Film Music: The Role of Emotion in Episodic Long-Term Memory for Music.” *BMC Neuroscience* 9.1 (2008): 9–48.

Forkel, Johann Nikolaus. *Johann Sebastian Bach: His Life, Art, and Work* (1802), trans. Charles Sanford Terry. London: Constable and Company, 1920.

Helmholtz, Hermann von. *On The Sensations of Tone* (1863), trans. Alexander John Ellis. New York: Dover Publications, 1954.

Heerema, Esther. “How Working Memory is Affected by Alzheimer’s Disease.” *Verywellhealth.com* (accessed October 14, 2023).

Henoch, Miriam A, and Kris Checky. “Ear Canal Resonance as a Risk Factor in Music-Induced Hearing Loss.” *Medical Problems of Performing Artists* 14.3 (1999): 103–106.

Jäncke, Lutz. “Music, Memory and Emotion.” *Journal of Biology* 7.21 (2008); 1–5.

Kardous, Chuck, Christa L. Themann, Thais C. Morata, and Gregory Lotz. “Understanding Noise Exposure Limits: Occupational vs. General Environmental Noise.” *Centers for Disease Control and Prevention* (blog accessed October 11, 2023.

Lhevinne, Josef. *Basic Principles of Pianoforte Playing*. New York: Dover Publications, 1972.

Liu, Fangfang, Shan Jiang, Jian Kang, Yue Wu, Da Yang, Qi Meng, and Chaowei Wang. “On the Definition of Noise.” *Humanities and Social Sciences Communications* 9.1 (2022): 8.

Maess, Burkhard, Stefan Koelsch, Thomas C. Gunter, and Angela D. Friederici. “Musical Syntax Is Processed in Broca’s Area: An MEG Study.” *Nature Neuroscience* 4.5 (2001): 540–545.

Malloch, Stephen, and Colwyn Trevarthen. “The Human Nature of Music.” *Frontiers in Psychology* 9.1680 (2018): 121.

Malloch, Stephen, Colwyn Trevarthen, and Ellen Dissanayake. *Communicative Musicality: Exploring the Basis of Human Companionship*. New York: Oxford University Press, 2009.

Manley, Geoffrey A. “Comparative Auditory Neuroscience: Understanding the Evolution and Function of Ears.” *Journal of the Association for Research in Otolaryngology* 18.1 (2016): 1–24.

McDermott, Josh H. “Auditory Preferences and Aesthetics.” In *Neuroscience of Preference and Choice*, edited by Raymond Dolan and Tali Sharot, 227–256. Cambridge: Academic Press, 2012.

McDermott, Josh H, and Andrew J Oxenham. “Music Perception, Pitch, and the Auditory System.” *Current Opinion in Neurobiology* 18.4 (2008): 452–463.

McDermott, Josh H, and Mark Hauser. “The Origins of Music: Innateness, Uniqueness, and Evolution. *Music Perception* 23.1 (2005): 49.

Merker, Bjorn. “When Extravagance Impresses: Recasting Esthetics in Evolutionary Terms.” *The Oxford Handbook of Music and the Brain*, ed. Michael Thaut and Donald A. Hodges, 67–78. Oxford: Oxford University Press, 2019.

Merriam-Webster.com (online dictionary https://www.merriam-webster.com).

"Musical Perception: Nature or Nurture?" *Neuroscience New*s (webpage accessed October 10, 2019).

Oxenham, Andrew J. "Pitch Perception and Auditory Stream Segregation: Implications for Hearing Loss and Cochlear Implants." *Trends in Amplification* 12.4 (2008): 316–31.

Peretz, Isabelle, Julie Ayotte, Robert J. Zatorre, Jacques Mehler, Pierre Ahad, Virginia B. Penhune, and Benoît Jutras. "Congenital Amusia: A Disorder of Fine-Grained Pitch Discrimination." *Neuron* 33 (2002): 188.

Pressnitzer, Daniel, Mark Sayles, Christophe Micheyl, and Ian M. Winter. "Perceptual Organization of Sound Begins in the Auditory Periphery." *Current Biology* 18.15 (2008): 1124–1128.

Ruwet, Nicolas. "Contradictions within the Serial Language." *Die Reihe* 6 (1964): 65–76.

Saldanha, E. L., and John F. Corso. "Timbre Cues and the Identification of Musical Instruments." *The Journal of the Acoustical Society of America* 36.11 (1964): 2021–2026.

Seashore, Carl E. "The Psychology of Music. VI. The Vibrato: (1) What Is It?" *Music Educators Journal* 23.4 (1937): 30–33.

Staecker, Hinrich, and Jennifer Thompson. "Central Auditory System, Anatomy." In *Encyclopedia of Otolaryngology, Head and Neck Surgery*, 376–383. Berlin: Springer, 2013.

Stiles, James. "The Decline of Serialism and the New Romanticism: Control and Chance in the New Music." *College Music Symposium* 19.1 (1979): 94–102.

Sundberg. "Acoustic and Psychoacoustic Aspects of Vocal Vibrato." *Speech, Music and Hearing Quarterly Progress and Status Report* (1994)*:* 45–68.

Terhardt, Ernst. "On the Perception of Periodic Sound Fluctuations (Roughness),"*Acta Acustica United with Acustica* 30.4 (1974): 201.

Trehub, Sandra, Glenn Schellenberg, and David Hill. "The Origins of Music Perception and Cognition: A Developmental Perspective." *Perception and Cognition of Music*, ed. Irène Deliège and John A. Sloboda, New York: Psychology Press (2015): 103-128.

Winkler, István, and Nelson Cowan. "From Sensory to Long-Term Memory." *Experimental Psychology* 52.1 (2005): 3–20.