The Bird's Ear View: Audification for the Spectral Analysis of Heliospheric Time Series Data

by

Robert L. Alexander

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Design Science)
in the University of Michigan
2015

Doctoral Committee:

Associate Professor M. Sile O'Modhrain, Co-Chair
Professor Thomas H. Zurbuchen, Co-Chair
Assistant Research Scientist Jason A. Gilbert
Professor Panos Y. Papalambros
Professor Emerita Mary H. Simoni

DEDICATION

For my mother,
whose light and love fill the sky.
Thank you for always encouraging me
to follow my dreams.
This song is for you.

ACKNOWLEDGEMENTS

The work presented in this dissertation is the final product of an incredible learning experience, and would not have been possible without the mentorship of the Design Science faculty and my advisory committee. I am incredibly grateful to my advisors from the School of Music Theatre and Dance, the department of Oceanic Atmospheric and Space Sciences, and the Design Science program for creating and fostering such a vibrant interdisciplinary learning environment. Over the course of this dissertation I faced many difficult and unforeseeable circumstances, and your support was unwavering.

First I would like to thank Sile O'Modhrain. Through your guidance my knowledge of experimental design has grown immeasurably, to the extent that I have trouble recognizing work I produced several years ago. Thank you for seeing my potential as a researcher and for continually holding my work to the highest standards. This dissertation would not have been possible without your wisdom and support, and I am incredibly grateful that you found time in your schedule every semester to hold weekly meetings. I enjoyed the many moments of laughter sprinkled throughout our intellectual discourse.

Before Thomas Zurbuchen heard the term "sonification" he had a hunch that listening to heliospheric data may provide new insight—this project took form when that spark of an idea met his entrepreneurial spirit. Thomas: as an advisor and mentor you set the bar high, and your commitment to the learning process has positively impacted the lives of countless students, myself being one of them. Thank you for going all-in.

I was incredibly fortunate to have a committee comprised of the finest researchers and distinguished faculty members. I would like to recognize Jason Gilbert for his tireless commitment to this dissertation. His support of the Rackham summer research proposal and the UROP mentorship made an invaluable contribution to my experience at the

University of Michigan. I am grateful to Panos Papalambros for building such an excellent interdisciplinary design community at the University of Michigan. I would like to thank Mary Simoni for connecting me with the Design Science program and for playing a crucial role in my early development as a researcher and critical thinker.

Diann Brei and Colleen Seifert have provided incredible guidance as I've grown as a researcher, and on behalf of all the Design Science students I would like to thank them for creating such an amazingly supportive and engaging learning environment. I've had the honor of passing through the program with a cohort of incredibly talented scholars. I would like to recognize all the support of the Design Science student body, past (Amir, Namwoo, Anna, Soodeh, Jihye, and Elliott) and present (Clover, Alex, Vineet, Melody, Matt, and Michael).

When I first met with the Solar and Heliospheric Research Group (SHRG) I had no formal understanding of the sun and Heliosphere, however, I was actively encouraged to ask questions and I am thankful that they were always well received. I would like to recognize Enrico Landi for tenaciously pursuing the investigation of Carbon Charge states. Thanks also to the members of the SHRG for your incredible contribution to this project: Thomas, Jason, Sue, Justin, Len, Jim, Jon, Jacob, Gina, Micah, Aleida, Mark, Pat, Paul and Liang.

The research scientists in Code 672 at NASA Goddard provided a remarkably high level of support for this project. I would like to thank Aaron Roberts for his guidance as a mentor, and for his valuable insights as the investigation unfolded. I would also like to recognize the contribution of Rob Wicks—the success of this research is due in large part to Rob's commitment to the project, and his eagerness to push across the boundaries of traditional analysis techniques. I am also grateful to Lan Jian for devoting a large portion of her time to the analysis of several data sets used in this dissertation. Thanks also to Brent, Bobby, Lynn, Mel, Georgia, Adolfo, Pablo, Chuck, Shing, Larry, Adam, and Michael for actively engaging with this research. I would also like to recognize the support of the interns at NASA Goddard, and the camaraderie of Ashley Jones and Dan Chuchawat.

A deepest heartfelt thanks is extended to my family. To my father, for always supporting me to follow my passion and for reminding me to always stand tall in the

arena. To my sister, Amanda, for being an incredible role model and a powerful force of good in this world—I‘ve always looked up to you, and words can’t express how grateful I am for your friendship and support. To my Mother, who will always live in my heart—every day I reflect upon how thankful I am to have been raised in such a loving environment. This dissertation is dedicated to her.

I would also like to extend my deepest appreciation for all the members of the Alexander and Mckay families, you are all incredible in your own way and I’m so honored to call you my family. Thanks also to my amazing friends for all their support, I admire you all and couldn’t have done this without you—I’ll personally express my gratitude the next time we cross paths.

I am grateful to all the participants who took part in the numerous studies presented in this dissertation, thank you for volunteering several hours of your time. I would also like to thank Linda Kendall and Eleanor Schmitt of the University of Michigan Design Lab for providing an excellent collaborative research space, and Eric Maslowski and Stephanie O’Malley from the 3D Lab for their support in creating incredible instructional/tutorial videos. I would also like to recognize Yohei Kanehara for contributions made to the audification code-base over the course of the UROP project.

I would like to thank the members of the ICAD community for welcoming me openly, and to Kelly Snook for facilitating so many valuable connections. I would also like to recognize the contributions made by co-authors on several papers, as well as the helpful feedback provided by many individuals in the drafting this document, including Mark Ballora, James Carson, Daphna Raz, Ajay Patel, Peter Raymond, Amanda Alexander, and Rachel Weinstein.

This work was supported by the NASA JPFP Fellowship program, the University of Michigan Solar and Heliospheric Research group, and the University of Michigan Design Science Program.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

ABSTRACT

The sciences are inundated with a tremendous volume of data, and the analysis of rapidly expanding data archives presents a persistent challenge. Previous research in the field of data sonification suggests that auditory display may serve a valuable function in the analysis of complex data sets. This dissertation uses the heliospheric sciences as a case study to empirically evaluate the use of audification (a specific form of sonification) for the spectral analysis of large time series. Three primary research questions guide this investigation, the first of which addresses the comparative capabilities of auditory and visual analysis methods in applied analysis tasks. A number of controlled within-subject studies revealed a strong correlation between auditory and visual observations, and demonstrated that auditory analysis provided a heightened sensitivity and accuracy in the detection of spectral features. The second research question addresses the capability of audification methods to reveal features that may be overlooked through visual analysis of spectrograms. A number of open-ended analysis tasks quantitatively demonstrated that participants using audification regularly discovered a greater percentage of embedded phenomena such as low-frequency wave storms. In addition, four case studies document collaborative research initiatives in which audification contributed to the acquisition of new domain-specific knowledge. The final question explores the potential benefits of audification when introduced into the workflow of a research scientist. A case study is presented in which a heliophysicist incorporated audification into their working practice, and the "Think-Aloud" protocol is applied to gain a sense for how audification augmented the researcher's analytical abilities. Auditory observations are demonstrated to make significant contributions to ongoing research, including the detection of previously unidentified equipment-induced artifacts. This dissertation provides three primary contributions to the field: 1) an increased understanding of the comparative capabilities of auditory and visual analysis methods, 2) a methodological framework for conducting audification that may be transferred across scientific domains, and 3) a set of

well-documented cases in which audification was applied to extract new knowledge from existing data archives. Collectively, this work presents a "bird's ear view" afforded by audification methods—a macro understanding of time series data that preserves micro-level detail.

Chapter I

Introduction

The sciences are inundated with a tremendous volume of raw data, and effectively navigating and analyzing growing data archives is a persistent challenge for the scientific community. At present, researchers typically extract knowledge from raw data by using various data visualization techniques. This reliance on visualization persists even though the benefits of auditory and multimodal displays have been well established within a growing body of literature in the fields of human computer interaction (HCI), psychology, interaction design, and data sonification. This last field, data sonification, is a highly interdisciplinary research area that explores "the use of non-speech audio to convey information" [1]. "Sonification" is an umbrella-term that encapsulates a wide variety of techniques for translating data into sound. A relatively transparent form of data sonification, known as audification, involves the direct presentation of frequency spectra through auditory means (while sonification involves a process of mapping). This technique has been documented as a potentially effective means for navigating and analyzing large data archives, however, empirical investigation into the usefulness of audification is largely lacking, and successful applications within the domain sciences are limited to a small handful of cases.

The body of work presented in this dissertation investigates the effectiveness of audification as a method for data mining, feature recognition, and knowledge extraction. Specifically, this dissertation uses data from the heliospheric sciences as a platform for formally developing, deploying, and evaluating audification methods. The overwhelming abundance of data produced by sun-observing instruments offers an excellent case study for this investigation—satellites maintained by the National Aeronautics and Space Administration (NASA) collectively produced an average 1.5 Terabytes (TB) of data per

day in 2004 [2], and the Solar Dynamics Observatory (SDO), launched in 2010, now single-handedly sustains a data stream of roughly 1.6 TB per day [3]. This torrential flow of information shows no sign of waning—the Daniel K. Inouye Solar Telescope, slated to come online in 2019, is projected to sustain data production rates of 10–20 TB per day, with possible peaks of up to 50 TB [4]. New data mining techniques are necessary in order to fully engage the massive archives produced by these instruments.

1.1. Traditional Analysis Methods

In a widely accepted definition, Hand (2001) describes the practice of data mining as "the analysis of (often large) observational data sets to find unexpected relationships and to summarize the data in novel ways that are both understandable and useful to the data owner" [5]. The data mining process draws common analysis techniques from an array of fields including statistics, machine learning, pattern recognition, and applied mathematics [6]. The exponential rise of raw computational power (i.e., related to Moore's Law) paints a bright future for scientific data analysis facilitated by automated analysis processes. However, data mining is a practice that, to be successful, often requires the active involvement of a human operator at every stage of the process [7]. Decisions based on informed expert assessments are essential in both confirmatory and exploratory investigations, and effective data mining should be considered a "discipline that must be mastered" [8]. At present, data visualization techniques provide the primary means through which scientific data (and the results of computational processes) are rendered accessible to human operators.

There are numerous ways to represent data visually. Line graphs allow researchers to plot multiple variables within a single visual space and are ideal for time series data. Other types of graphic representation, such as histograms, scatter plots, and bar graphs allow researchers to visually analyze data for high-level features or patterns [9]. Visualization tools utilize parameters such as color, size, and shape to render data sets of moderate complexity. However, early studies conducted by Pollack [10], Fidell [11], Loveless [12], Yeung [13], and Bly [14] suggest that visual analysis methods may be greatly enhanced by supplementing visualization with auditory display—and in some

instances the use of auditory analysis techniques may be preferable as the primary method of survey (this body of research is reviewed in Chapter II).

1.2. Sonification and Audification

The *Sonification Handbook*, an open-access resource published in 2012, established an important set of best practices, laying a much-needed foundation for ongoing research initiatives [15]. The sonification community now recognizes a set of formal techniques for data sonification, including *parameter mapping, auditory icons, earcons, auditory graphs, visual substitution, model-based sonification*, and *audification*. Audification is a translational process through which data samples are isomorphically mapped to the amplitude values of consecutive audio samples. This is the most direct form of sonification, as all samples from the original data set are preserved (see **Fig. 1**).

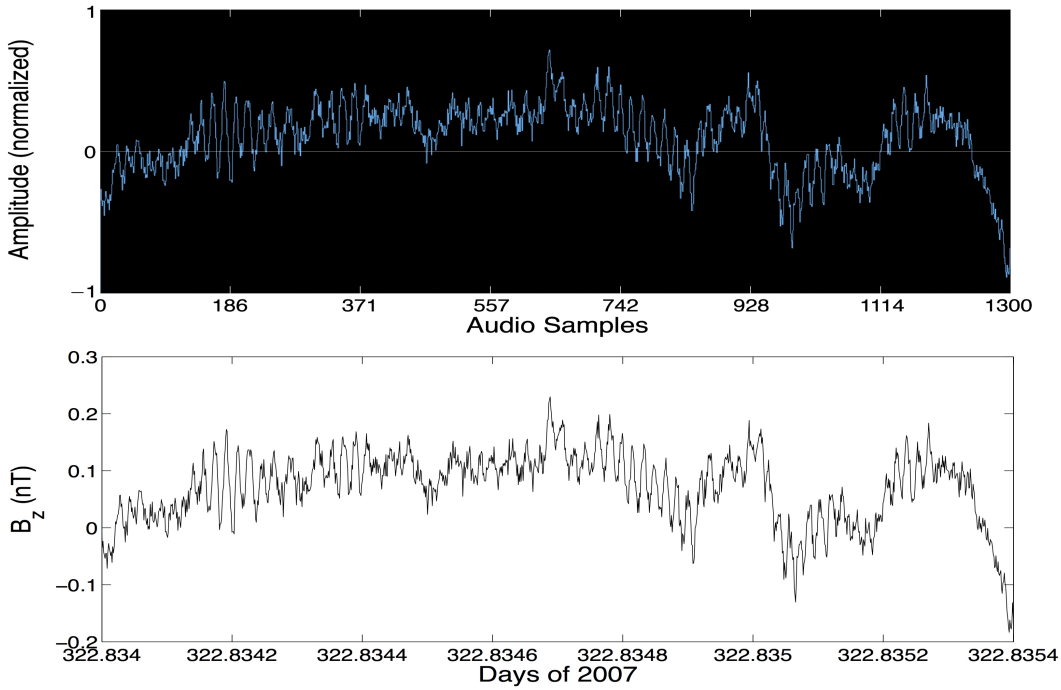


Figure 1. Thirteen-hundred samples of high-resolution magnetometer data from the *Wind* satellite (*z* component). The audio waveform rendered in iZotope Rx (top) and the time series rendered in Matlab (bottom) are identical, as no data samples are lost in the conversion process (reprinted from Alexander et al. 2014, fig. 1).

When an audified data set is played over speakers or headphones, spectral features within the original data are translated as timbral components in the resulting audio. This process exploits the innate pattern-matching abilities and high temporal

resolution of the human auditory system. As the standard rate of sound file playback is 44,100 samples per second, audification allows for the examination of one million data samples in less than 23 seconds. For this reason audification provides a promising method for evaluating large time series data sets, particularly for exploratory purposes.

1.3. Scope

This dissertation extends existing research that investigates the use of audification as a means for analyzing time series data. Audification has been successfully applied in the field of geo-seismology for purposes of data survey and feature identification, where Hayward (1994) suggests it may reduce the likelihood that important features will be overlooked (such as equipment induced noise) [16]. Auditory seismograms have been qualitatively evaluated within the International Community for Auditory Display (ICAD), though user studies are necessary in order to quantify the results of this work [17, 18]. Pauletto and Hunt (2005) uncovered a strong correlation between assessments made through audification and visual analysis of spectrogram displays rendered from complex time series data [19]. This investigation established a foothold for audification within scientific culture that was previously dominated by visual analysis methods. A number of informal experiments have been conducted to evaluate the effectiveness of audification techniques for identifying patterns in stock market data [20]—Worrall (2010) for example noted that audification was appealing for this purpose, though the investigation was limited due to the use of short data examples (a set of controlled experiments was suggested in order to determine the nature of features that may be perceived within audified data sets). Most recently, a parameter mapping approach used for exploring space physics data through sound revealed that scientists were highly sensitive to the presence of subtle features in the data (this work is reviewed in Chapter II) [21].

This dissertation explicitly focuses on the use of audification as a method for data mining and analysis. Hayward's assertion that audification may reveal features that would have otherwise been missed is a central focus of this investigation. As its point of departure, this dissertation directly extends the work conducted by Pauletto and Hunt (2005) and is framed within a domain-specific investigation into the comparative capabilities of auditory and visual analysis methods as applied by heliospheric scientists.

To date, data audification has been adopted into scientific workflows on a relatively *ad hoc* basis. For example, audification successfully led to the classification of broadband noise bursts that were produced as the Voyager 2 spacecraft collided with Saturn's ring particles [22], however, in this case the application of audification is not well documented.

A wide variety of phenomena documented within the canon of space physics were originally detected through auditory means, including *whistlers* [23], *chorus* [24], *hiss* [25], and *lion roars* [26]—these phenomena owe their descriptive names to auditory observations. However, while the historical contribution of audification methods provides anecdotal support for the claim that exploratory audification can reveal new features within the heliospheric data sets, in order for audification to gain widespread acceptance within domain sciences a great deal of research needs to be completed on the relative effectiveness of the auditory modality when applied to complex data analysis tasks.

1.4. Research Questions

The work presented here systematically explores the application of audification methods for the analysis of large time series data sets, focused primarily on case studies conducted with heliophysics research scientists at the University of Michigan and NASA Goddard Space Flight Center. In addition, this research forms the foundation for the systematic integration of an auditory analysis toolset in a range of applications—a multifaceted design challenge that necessitates a firm understanding of the perceptual limitations of the auditory system. Case studies presented here investigate the use of audification as an effective means for surveying large quantities of time series data and gaining what may be described as a "bird's ear view"—a perspective that enables the evaluation of macro-scale structures while simultaneously providing the nuanced details of small-scale spectral features. This research is guided by the following primary questions:

Q1: How do auditory analysis capabilities compare with visual analysis methods in the evaluation of large time series data sets?

Q2: Can audification reveal features that may be overlooked through visual analysis of spectrogram displays?

Q3: What benefits can audification provide when introduced into the workflow of research scientists dealing with large time series data sets?

The first inquiry is addressed through a number of tightly controlled studies in which participants were asked to visually and auditorily assess a variety of time series produced by sun-observing spacecraft. The first two studies incorporate parameterized artificially generated stimuli such that the nature of features observed through auditory and visual analysis methods may be quantitatively assessed. The remaining three studies focus explicitly on the classification and identification of various naturally occurring phenomena within heliospheric time series (e.g., low frequency wave storm events). In the second, third, and fifth studies, features within the data are assessed with respect to the number of positive identifications that were made visually and auditorily—providing a quantitative measure for **Q2**. This dissertation also addresses **Q2** in a number of well-documented case studies in which an audification specialist worked alongside heliospheric research scientists in the analysis of space mission data archives. The investigation of **Q3** employs a technique known as the Think-Aloud protocol (an approach that is commonly applied in the fields of human-computer interaction and software design) to gain a sense for how auditory analysis methods contribute to the analysis process of a research scientist who has some level of familiarity working with audified data. The majority of studies incorporate some discussion in which results are interpreted in terms of the potential benefits that audification may provide to traditional analysis workflows.

1.5. Primary Contribution

This research provides three primary contributions to the field of data audification. First, knowledge gleaned through five empirical studies provides a new understanding of the effectiveness of auditory and multimodal display as a means for the spectral analysis of long time series. Second, this work presents a methodology through which researchers can adopt audification techniques into their existing working

practices.[1] Finally, this dissertation contains a set of well-documented case studies in which audification played a crucial role in producing new domain-specific knowledge.

1.6. Thesis Outline

This body of work is presented within five chapters that provide the primary qualitative and quantitative contributions of this research (Chapters III–VII). This contribution is book-ended by a review chapter that provides introductory context and a final chapter that summarizes the contribution of the research and discusses future directions. **Figure 2** provides a visual-reference that links the research questions with the primary chapters and sections in which they are addressed.

Chapter II reviews the functionality of the human auditory system before investigating the psychoacoustic principals that play a role in the auditory data analysis process. The studies presented in **Chapter III** directly compare auditory and visual analysis capabilities in applied analysis tasks in order to establish baseline metrics for auditory performance. **Chapter IV** contains a case study in which a heliophysics research scientist was introduced to audification and his evolving interaction was documented through the Think-Aloud protocol; this chapter also outlines several novel audification use cases. **Chapter V** outlines several case studies of specific applications of the proposed methodology, documenting work conducted with the Solar and Heliospheric Research Group (SHRG) in which audification played a crucial role in identifying subtle spectral cues in carbon charge-state data from the Advanced Composition Explorer (ACE) spacecraft, and several interactions with research scientists at NASA GSFC in which audification was successfully applied in the identification of novel spectral features in large time series data sets. Based on knowledge acquired through the interactions documented in the fifth chapter, **Chapter VI** establishes a methodological framework for conducting auditory analysis through audification, focusing on the technical and practical aspects of audification as a method for feature identification. **Chapter VII** empirically tests the method established in Chapter VI, presenting a within-subjects study in which twenty heliophysics research scientists were asked to visually and

[1] The twelfth chapter of the sonification handbook provides an excellent introductory resource for individuals who are interested in learning the practice of data audification [27], however, this dissertation provides a much-needed framework for audification as applied toward exploratory and confirmatory analysis.

auditorily identify low frequency wave storm (LFWS) events in data gathered by the *Wind* spacecraft. The results of this study are investigated through metrics derived from binary classification theory. **Chapter VIII** reviews the main contributions of the thesis, situates the outcomes of this work within the context of sonification research that has been conducted to date, and suggests directions for future research.

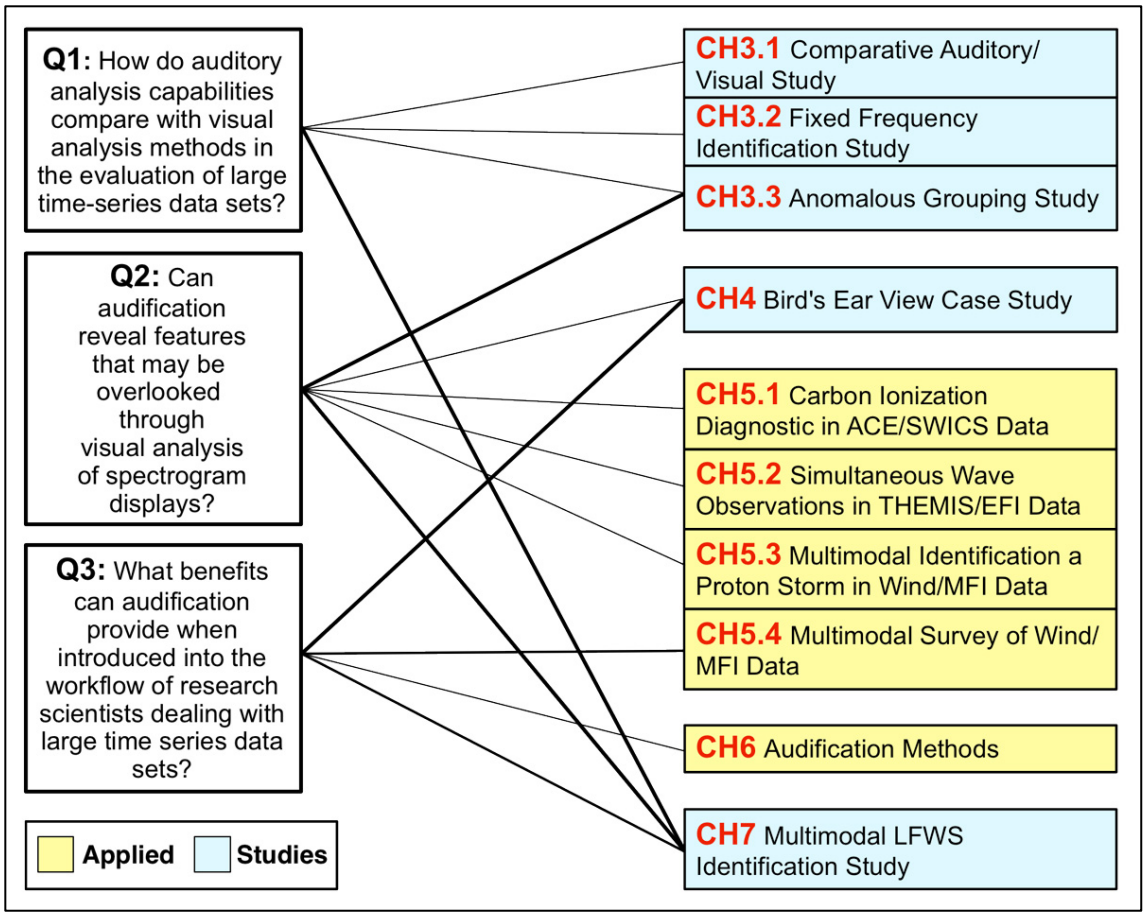


Figure 2. A map outlining the primary research questions and the chapters/sections in which they are addressed. A line's thickness indicates the size of the contribution.

CHAPTER II

The Science of Listening

The branch of science that concerns itself with human auditory perception is known as psychoacoustics. This chapter will begin by situating the auditory data analysis process within a larger perceptual framework—fusing a bottom-up understanding of auditory physiology with a top-down understanding of auditory stream segregation and cognition. Far from a "black box" system, human audition has well-known limitations that should be considered in the evaluation of audified data sets. The remainder of this chapter will document the origins of various sonification techniques along with ideal use cases, ending with a review of the history of audification in the space sciences.

2.1. Fundamentals of Auditory Perception and Psychoacoustics

Acoustics is the branch of physics that deals with the generation, transmission, and absorption of mechanical vibrations (of any frequency or amplitude) [28]. The range of vibrations that fall between 20 Hz and 20 kHz can be referred to as "sound," as this is the full range of auditory perception in healthy individuals. Psychoacoustics is the branch of science that deals with the perception of sound—incorporating such disciplines as music, psychology, physics, engineering, audiology, physiology, and linguistics [29]. This section will begin with a brief review of the functionality of the human auditory system, including well-known perceptual biases, sensitivities, and limitations.

2.1.1. **The Physiology of Audition**

Individuals must parse a complex set of stimuli while navigating everyday environments. The acoustic startle response (ASR) is thought to play a crucial protective role in preparing for fight/flight [30], and average reaction time in the detection of auditory stimuli (140–160 ms) has been found to be faster than for that of visual stimuli

(180–200 ms) [31, 32]. The temporal resolution of the ear is remarkably high compared to that of the eye—the standard rate of sound file playback is 44,100 Hz while the eye perceives smooth visual motion at a frame rate of 60 Hz. This says nothing of the relative bandwidth of information carried by the two sensory modalities, or the relative size of the auditory and visual cortex, but it does speak to the unique functionality of each system. The auditory system is specifically well suited for the assessment of phenomena that unfold at a high temporal resolution. This is understandable from an evolutionary standpoint, as the avoidance of an unseen predator is a critical survival skill, and in this respect reaction time is crucial. But exactly what happens in that moment when a stick breaks in the underbrush, or a car horn blares on the highway?

First, objects in the environment generate a wave of compression and rarefaction that travels through a medium such as air or water, which carries these perturbations [32]. These waves arrive at the **outer ear**, travel through the auditory canal and come into contact with the tympanic membrane (eardrum), causing it to vibrate. This vibration is passed through a set of three small bones known as the ossicles (commonly referred to as the hammer, anvil, and stirrup—the correct anatomical names are the malleus, incus, and stapes). These bones constitute the **middle ear**.[2] The stapes connects directly to a small membrane known as the vestibular window (i.e., oval window) which transfers the vibrational energy into two fluid filled tubes contained within a spiral-shaped cavity known as the cochlea. The acoustic energy has now arrived at the **inner ear**. A structure within the cochlea known as the basilar membrane separates these tubes and extends from the base (the region connected to the vestibular window) to the apex (the far inner tip). The basilar membrane has a distinct vibrational response to sound, as a wave of displacement travels from the base to the apex [34].[3] This wave (known as the traveling wave) increases in magnitude as it propagates until it abruptly stops at a specific location. The location where this traveling wave reaches peak intensity is directly correlated to the frequency of the sound [34, 35]. In this way, frequency is tonotopically distributed across the basilar membrane and the function of the cochlea is similar to that of a Fast Fourier

[2] Muscles between the ossicles contract when exposed to loud sound in order to minimize potential damage to the delicate structures of the inner ear—an action known as the acoustic, or stapedius reflex [33].
[3] Much of what is known about the functionality of the basilar membrane and the cochlea has origins in the work of George von Békésy who won the Nobel prize in Medicine for this research in 1961.

Transform (FFT) [36].[4] This understanding of cochlear function is known as "place theory," and is commonly attributed to Hermann Helmholtz [36]. **Figure 3** situates a simplified graphical representation of the basilar membrane alongside a spectrogram display[5] of the words "hello world" spoken into a microphone—this provides a visual depiction of the evolving activation pattern of the membrane in response to acoustic energy.

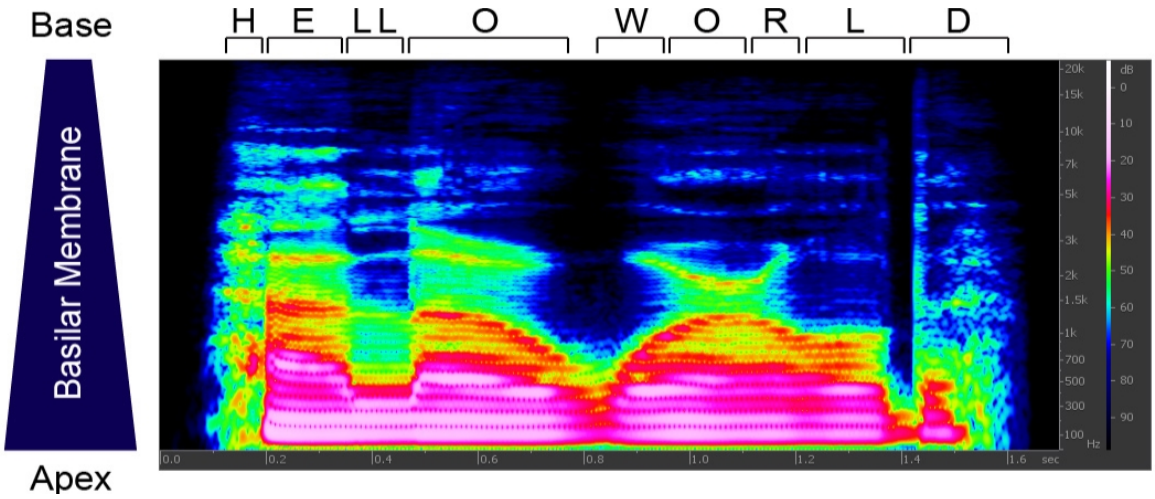


Figure 3. A spectrogram display of the words "hello world" spoken into a microphone (displayed in iZotope Rx) alongside a representation of the basilar membrane. The spectral power of each letter is displaced from base to apex accordingly.

A number of hair cells are displaced along the basilar membrane—about 12,000 outer hair cells, each with about 140 hairs (known as cilia), and approximately 3,500 inner hair cells, each with about 40 cilia [37]. Movement of the basilar membrane results in a shearing motion that causes the tops of these hair cells to be displaced.[6] These inner hair cells, then, transduce this displacement into neural activity [36]. This activity is projected by axons in the auditory nerve into the nervous system in a systematic way

[4] For frequencies above 500 Hz, the distribution of intensity along the basilar membrane is related to the logarithm of the frequency of the sound. For a more complete review of this functionality see Moore (2005) pp. 23–32.
[5] All spectrogram displays provided in this dissertation will display time along the *x*-axis, progressing from left to right, and frequency along the y-axis ascending from low to high.
[6] These hair cells are contained within a region known as the "organ of corti," and the tips of these cells rest against the tectorial membrane. It is the shearing motion against this membrane that causes displacement; see Moore (2005) pp. 32–33.

such that the tonotopic representation of frequency is preserved in the ascending nervous system [38, 39] and auditory cortex [40].

2.1.2. **Loudness**

The intensity of a sound is traditionally measured in *decibels*,[7] a unit that provides a logarithmic measure of sound pressure level (SPL). An increase of approximately 3 dB corresponds to a doubling of subjective loudness, while a 10 dB increase corresponds to a factor of 10. The most intense sound that one can be endure without immediate hearing loss is approximately 120 dB louder than the quietest perceptible sound—a dynamic range spanning approximately 12 orders of magnitude (i.e., 10^{12}) [36]. Measures of subjective loudness are traditionally provided in dB(A), a weighted measure defined in 1936 in the *American Standards for Sound Level Meters* [41]. **Table 1** includes a list of common sounds and their associated dB(A) level for reference [42, 43].

Table 1. Loudness levels associated with familiar sounds, adapted from Rabinowitz (2000) and Stevens (1959).

Loudness dB(A)	Sound
150 - 170	Gunshot
140 - 150	Jet takeoff
140	Threshold of pain
120	Threshold of discomfort
110 - 120	Rock concert / Chain saw
100	Diesel locomotive
94	Hearing loss after 1 hour exposure
90	Motorcycle / Lawnmower
80	Heavy traffic
60 - 65	Conversation
50	Quiet room
30 - 40	Whisper
0	Threshold of hearing

The *A-weighted* scale takes into consideration the variation of perceived loudness as a function of frequency in order to provide a subjectively balanced measure of environmental noise [44]. These measures were established by the work of Fletcher and

[7] The prefix deci- indicates one tenth and the suffix –bel is a unit coined in honor of Alexander Graham Bell.

Munson (1933), who, through a controlled study, measured absolute versus perceived intensity across a wide range of frequencies [45]. This measure of perceived intensity can be graphically depicted (see **Figure 4**) and the resulting contours are commonly referred to as "equal loudness curves" [45]. Tracing along any given curve one can determine how loud a pure-tone must be played in order to match the subjective loudness of a pure tone played at a higher or lower frequency. Observing **Figure 4**—a 40 Hz tone played at 80 dB will perceptually appear to be equal in loudness as a 2 kHz tone played at 60 dB, despite the fact that acoustically speaking the 40 Hz tone carries 100 times the power. A quick visual scan will reveal that the peak sensitivity of the human auditory system occurs between 2–5 kHz, a range crucial in the detection of human speech.

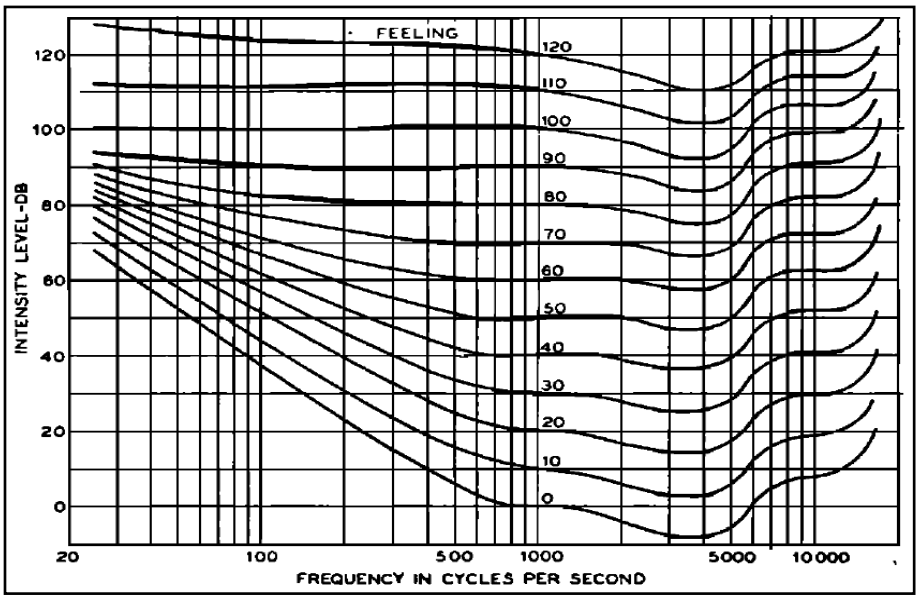


Figure 4 Loudness Level Contours reprinted from Fletcher and Munson (1933), fig. 4.

A large number of factors can play into the subjective loudness of a given stimulus, including the spectral bandwidth and temporal duration. For example, the perceived loudness of a cluster of pure-tones will remain constant until the spacing between the tones is increased past a "critical band," after which perceived loudness will increase. This same increase will occur when the bandwidth of band-limited noise is

enlarged (**Figure 5** in the following section depicts the critical bandwidth as a function of frequency).[8]

2.1.3. **Pitch**

The sensation of *pitch* is determined by the periodic rate at which pressure fluctuations arrive at the eardrum [38]. This subjective measure could be defined as the perceived "highness" or "lowness" on a scale, and in the case of a pure sinusoidal waveform pitch is directly correlated with frequency [29]. For reference, **Table 2** provides the corresponding frequency and wavelength (cm) of C played across the standard range of a musical keyboard. This information may be important while translating auditory observations into the original time scale of an audified data set.

Table 2. The frequency of C across 8 octaves, and the corresponding wavelength.

	Frequency (Hz)	Wavelength (cm)
C 1	32.7	1054.9
C 2	65.4	527.5
C 3	130.8	263.7
C 4	261.6	131.9
C 5	523.3	65.9
C 6	1046.5	33.0
C 7	2093.0	16.5
C 8	4186.0	8.2

The human ear is remarkably sensitive to subtle frequency shifts in periodic waveforms. The smallest perceptible change in pitch is defined as the Just Noticeable Difference (JND)—at 100 Hz the ear is capable of detecting a JND of approximately 1–2 Hz (1–2%), and at 1 kHz the JND increases to approximately 3–5 Hz (0.3–0.5%) (see **Figure 5**) [38, 46]. The JND provides a reasonable estimate for the error-margin that may be expected when the auditory observation of a fixed-frequency signal is reported. Our ability to resolve discrete pitches (and hence intervallic relationships between pitches) is

[8] For a thorough review of spectral and temporal factors that mediate subjective loudness see Zwicker et al. (1957) and Moore (2004) pp. 134-137.

finely tuned at low frequencies, while intervals increase in subjective magnitude at higher frequencies [47].[9]

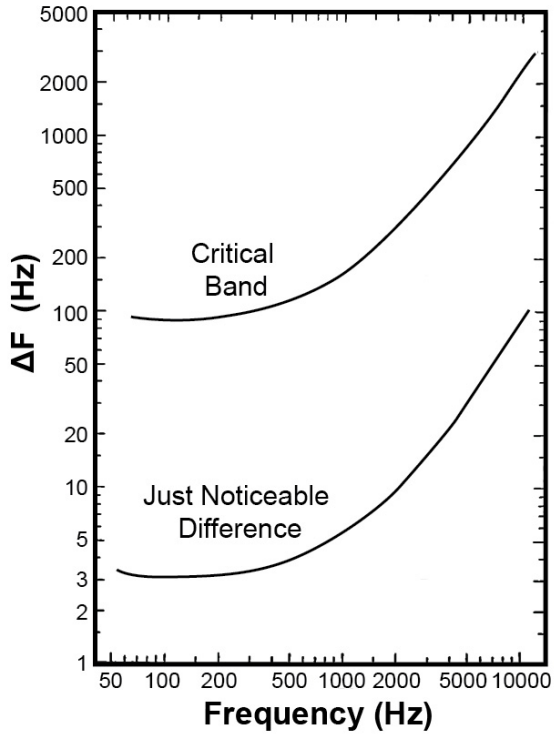


Figure 5. Critical band and Just Noticeable Difference as a function of frequency (adapted from Zwicker et al. 1957, fig. 12).

2.1.4. **Spatialization**

Just as the two eyes are able to process light from the surrounding environment to form the unified perception of height, width, and depth (stereoscopic vision), so too are the ears able to process subtle timing and pressure changes in the air to give rise to the auditory perception of three-dimensional space (stereophonic audition). The effortless nature with which human beings are able to survey their environment belies the underlying complexity of the mechanisms that give rise to a seamless sensory experience.

[9] The Mel scale was developed to provide a quantitative metric for perceived pitch magnitude and the Bark scale provides an alternative mapping between critical-band width, and frequency (in this system, one bark corresponds to one critical band) [48].

The auditory system integrates a series of spectral and temporal cues to provide a nearly instantaneous encoding of an object's physical location. In the case of data audification, spatial cues may be exploited for analysis purposes when data parameters are played back across multiple audio channels. In exploring these cues, one may consider the sound field experienced by an individual snapping their finger within a relatively quiet space. With the hand lifted directly in front of the face (on the median plane bisecting the two ears) the sound of a snap will reach both ears at the same time and with the same relative intensity. When the hand is moved off to the left, the sound will subsequently arrive at the left ear slightly before the right, and with a slightly higher intensity—these phenomena are known as interaural intensity difference (IID) and interaural time difference (ITD) respectively [36].

A number of factors mediate IID including the "acoustic shadow" produced by the head, which has a larger impact on the localization of higher frequency sounds [49]. Low frequency sounds have a comparatively longer wavelength and hence are able to curve around the head through a process known as diffraction [36]. As a sound source moves closer to either ear it grows exponentially louder in accordance with the inverse square law,[10] and in close proximity the IID is pronounced at both low and high frequencies.

ITD will be non-existent when a sound is produced on the median plane and approximately 690 microseconds when a sound is produced 90° to one side. In the case of a sinusoidal tone, timing differences amount to an interaural phase difference (IPD) across the two ears—IPD provides a reliable spatial cue at low frequencies, and is more ambiguous at higher frequencies. In the case of pure sinusoidal waveforms, IID provides the dominant spatial cue for high frequencies, while ITD dominates low-frequency spatialization [36, 50].[11]

A transient sound occurring within an enclosed space will potentially reflect off any number of surfaces (e.g., walls) resulting in a cascade of echoes that arrive at the

[10] The intensity of a sound is inversely proportional to the square of the distance between the sound source to the receiver.

[11] In the case of a conical cylinder extending outward from the ear, any sound produced along the circumference of a bisected region of this cone would result in the same ITD, hence this cue provides some locational ambiguity. This "cone of confusion" has found to be resolved through movement of the head, see Hirsh (1971) and Moore (2005) pp. 248–249.

listener. A complex phenomenon known as the "precedence effect" is observed when these echoes are not perceived as separate reflections, but rather are perceptually fused with the direct sound to create the impression of a single localized sound source. This same effect is observed over headphones when a transient sound (such as a snap or a click) is played with an ITD of 1–5 ms, or in the case of complex sounds such as speech, an ITD of 1–40 ms. The effect is lost when the ITD is increased above these thresholds, in which case the delayed sound is perceived as a separate event [36, 51]. [12]

The structure of the outer ear plays a large role in encoding spatial perception. Front-back and up-down localization are largely determined by spectral cues provided by the pinna, a small flap of cartilage that directs sound into the ear canal [52]. It has been documented that individuals with unilateral hearing loss retain the ability to localize sound entirely based on these spectral cues [49]. The directionally dependent filter imparted by the head and the structure of the outer ear will vary from individual to individual—this effect can be quantified by comparing the spectrum of a sound in a free field environment to the spectrum that arrives at the eardrum. The ratio of these two spectra is known as the "head related transfer function" (HRTF) [36]. While the spatial field experienced within headphones is generally referred to as "lateralization," binaural encoding can be simulated with the application of an HRTF to create the illusion of localization within three-dimensional space [36, 53].

2.1.5. **Masking**

In an early experiment, A. M. Mayer documented that the presence of a low frequency "obliterated" the sensation of a second, higher frequency sound. In these experiments, the sound of a low organ pipe was found to "smother" that of a high tuning fork struck with considerable intensity, though it was observed that, "no sound, even when very intense, can diminish or obliterate the sensation of a concurrent sound which is lower in pitch." A second experiment found a similar effect in the simultaneous ticking of a large and small clock—in which case the lower frequencies emitted by the larger clock completely obscured the ticking of the smaller clock [54, 55].

[12] This effect is also broken when amplitude of the delayed sound is sufficiently increased.

This effect—now commonly referred to as "masking"—is defined by the *American Standards Association* as "The process by which the threshold of audibility for one sound is raised by the presence of another (masking) sound." This definition proceeds to specify dB as the unit for measuring this threshold [56]. In the case of audified data, it is possible that certain spectral features may subjectively overshadow the presence of others.[13] This masking effect has unique properties for pure tones, complex sounds, narrowband noise and broadband noise [29], which can be summarized as follows [29, 57, 58]:

Pure Tone Masking

1. Lower frequency tones provide an effective mask for higher frequency tones, but the reverse is not true.
2. The strength of the effect is inversely related to the distance in frequency between the tones, where widely separated tones will impart little or no masking.
3. A pure tone of higher intensity will effectively mask a wider range of frequencies.

Noise Masking

4. Narrowband noise provides an effective mask in a manner similar to pure tones, in which case the characteristics of pure tone masking **(1– 3)** generally hold true.
5. Broadband noise provides an effective mask for tones at all frequencies, with a direct linear correlation between the intensity of the masking effect and the intensity of the noise mask.

Nonsimultaneous Masking

6. Forward masking occurs when a sound that abruptly ends obscures the perception of a tone presented shortly thereafter (1–30 ms).
7. Backward masking is the reverse of **(6)**—generally occurring when the presence of a tone is obscured by a sound occurring 1–10 ms after. The effect of backward masking can be minimized or eliminated with training.

This effect can be considered as arising from the resonant response of the basilar membrane: low frequencies trigger a wave of excitation that travels upward toward the apex; this excitation pattern is asymmetrical—abruptly cutting off toward the apex, but extending toward the base (where the transduction of higher-frequencies occurs) [29].[14]

[13] The impact of a masking signal on auditory and visual feature-identification is explored in the third chapter.

[14] Masking also results from higher-level processes, as in the case of "central masking," in which a tone presented in one ear can mask the presence of a tone in the other ear. See Rossing (2002) p. 115.

Fletcher (1940) determined that in the case of band-limited noise, the "critical bandwidth" at which masking occurs is equal to the ratio of the intensity of the masked tone to the average intensity per cycle of the noise producing the masking [58]. Fastl and Zwicker (2007) provide the threshold for audibility of a sinusoidal tone when played with a 60 dB band-limited noise mask with a center frequency of 250, 1000, and 4000 Hz (with a bandwidth of 100, 160, and 700 Hz, respectively). The results of this work are adapted in **Figure 6,** which displays the threshold of audibility in a quiet space (which adheres to the equal-loudness contour established by Fletcher and Munson, 1933), the frequency of each 60 dB noise mask, and the region obscured by the mask (highlighted in blue) [57].

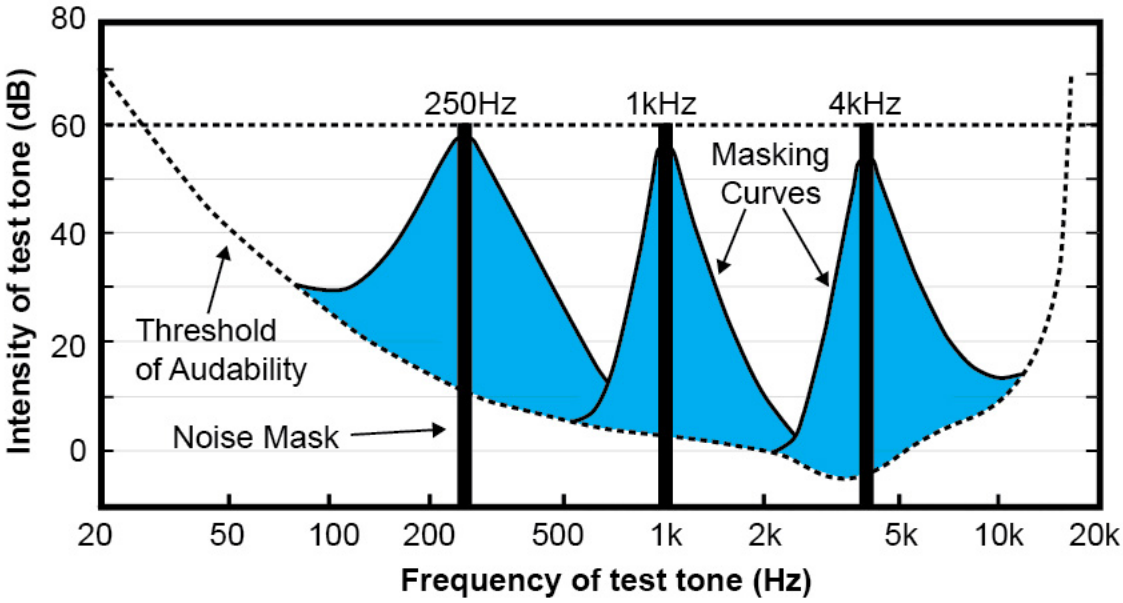


Figure 6. Level of test tone just masked by band-filtered noise with a level of 60 dB, and center frequencies of 0.25, 1, and 4 kHz. The broken curve is the threshold of audibility in the absence of a noise mask. The blue areas underneath the curves indicate frequency regions that will be masked (adapted from Zwicker, 2007).

2.1.6. **Fatigue and Hearing Loss**

The effects of long term listening on the human auditory system must be considered when conducting any extended auditory investigation. Exposure to intense sounds for a long duration of time can lead to temporary threshold shifts within the auditory system (as well as permanent threshold shifts in extreme cases). This temporary

threshold shift is the most common index for measuring auditory fatigue, which is defined here as "the application of a stimulus which is usually considerably in excess of that required to sustain the normal physiological response of the receptor" [59]. Major factors that impact the degree of shift include the intensity, duration, and frequency of the fatiguing stimulus [36]. Temporary threshold shifts can lead to a reduction in temporal summation, poor frequency resolution, abnormally rapid loudness growth, and poor speech discrimination [60]. Many professionals in the audio industry rely on their ears for "technical listening" (e.g., music producers and mastering engineers)—these individuals are keenly aware of the need to take incremental breaks in order to prevent ear fatigue [61].

Exposure to noise can result in irreversible hearing loss at sustained levels equal to an average SPL of 85 dB(A) or higher for an eight-hour period [42]. The required exposure time for potential damage to the auditory system is cut in half with each increase of 3 dB, hence the same result will be achieved from an exposure to a stimulus at 94 dB sustained for one hour, or 112 dB for one minute. This damage occurs when a sound is sufficiently loud to exert a shearing force on the cilia, damaging the cells, and potentially resulting in cell death [42].

2.1.7. **Timbre**

The definition of timbre provided by the American Standards Association is "that attribute of auditory sensation in terms of which a listener can judge that two sounds similarly presented and having the same loudness and same pitch are dissimilar" [56]. While this definition is apt in the comparison of two tonal instruments (timbre is what differentiates the sound of a trumpet from a violin, or a voice from a flute), timbre may also be considered as a descriptor for the unique distribution and evolution of spectral power in a single complex sound [62]. This evolutionary quality is important—the timbre of a cello may initially be "dark" while growing "brighter" over time, and we may consider the fact that the timbre of a recorded piano resembles that of an accordion when played in reverse [36].

A French horn shares many common spectral components with a saxophone, but the evolving "amplitude envelope" of each serves to readily differentiate the two [29].

We can describe this envelope as having an **attack** (the gradual or abrupt nature of a sound as it emerges from silence), **decay** (the fall to a steady state after this initial attack), **sustain** (the body of a sound as it perpetuates), and **release** (the relative smoothness with which a sound fades into silence) [63]. This unique amplitude contour plays a role in defining the timbre of a sound, and in the task of differentiating one sound from another—in this sense, timbre provides an excellent segue into the larger question of how the auditory system assigns "belongingness" to a given subset of auditory stimuli.

2.2. Auditory Scene Analysis

Beginning in the 1960's, the recognition of visual objects by computational methods has been a problem addressed by the field of computer vision. This task (of deconstructing images into their requisite parts) is known as scene analysis. The language and principles of *Auditory Scene Analysis* were introduced by Bregman in a 1978 paper [64] and later expanded in a full volume of the same name [65]. In lieu of visual *objects*, the task becomes one of classifying auditory *streams*—a proposed perceptual unit for a single auditory "happening." Auditory data analysis, then, may be considered a task in which the analyst separates meaningful signals from background noise through a process of *auditory stream segregation* [65]. Several chapters in this dissertation contain studies that assess the ability of participants to identify spectral features presented visually and auditorily, and in many instances results are interpreted within the framework of auditory scene analysis—for this reason an understanding of Bregman's framework is important. This section will introduce the principles of scene analysis, and examine the process of stream segregation through the lens of Gestalt psychology.

2.2.1. **Auditory Streaming**

Bregman defines an auditory stream as "a psychological organization whose function it is to mentally represent the acoustic activity of a single source over time" [66]. In approaching this definition it is helpful to return to visual perception, and consider how vision contributes to the mental representation of an object such as an apple—we perceive a number of distinct attributes such as color, texture and shape, and this set of attributes is essential in classifying types of apples, as well as segregating apples from various other objects. In defining any given percept as an auditory object (or an auditory

stream) we can again draw upon a number of attributes along the lines of timbre—the tone of a violin has characteristic "color" different from that of a cello, and the low resonance of the cello may be described as having a greater sense of "roundness." The principle of auditory streaming goes one step further in considering the psychological principles that allow one to differentiate between the sounds of two violins playing in unison, or to identify the sound of a single violin playing on a subway platform as a train pulls into the station. The latter task presents a complex scene analysis problem, but one that is not beyond the abilities of the auditory system.

In the mid 1950's a researcher at MIT, Colin Cherry, undertook a number of experiments to explore what he called the "cocktail party problem" [67]. These experiments investigated the various perceptual mechanisms that play a role in the process of selective attention, and the *cocktail party effect* is now a common descriptor for the ear's ability to attune to a single auditory stream among a number of potential distractors.[15] While the high-level processes that mediate speech perception are beyond the scope of this introduction, the complex frequency spectrum of the human voice provides an excellent case study for the formation of auditory streams.

2.2.2. **Stream Integration and Perceptual Fusion**

Stream integration provides a conceptual model for the process through which the auditory system assigns a subset of incoming sensory stimuli a sense of unified "belongingness" based on temporal and/or spectral cues [65]. A number of experiments in sound-synthesis were conducted in the 1980's that began to shed light on the facets of complex auditory streams that give rise to the perception of voice. The acoustic theory of human vocalization had previously established that vowel sounds are generated through unique resonances in the vocal tract (i.e., formants)—the sound of "e" for instance, is produced by strong spectral peaks at approximately 430, 1960, and 2720 Hz [68]. However, simply synthesizing pure tones at these frequencies does not result in a sound with a distinctly "human" quality. In order to explore this perceptual landscape, John

[15] This process is described by Cherry (1953) as "filtering," and the original experiments explored the important role of spatialization and probabilistic phrases.

Chowning[16] designed an experiment in which a soprano voice was digitally synthesized in three stages [70]:

Stage 1: A sinusoid was generated at the fundamental frequency.

Stage 2: Pure tones were added, appropriate to the harmonics of a sung vowel.

Stage 3: A mixture of random pitch variation and vibrato were added to the total signal.

If the sonority of a voice were achievable through the mere combination of spectral formants, then we would expect the vocal tonality to emerge in stage 2. However, it was reported that at this stage the harmonics did not cohere and were perceived as separate tones. Only in stage 3, when the vibrato and pitch variation were added, did the synthesized sound resemble that of a human voice—these added time-based elements gave rise to perceptual fusion [70]. It can be said that each of the tones was segregated as a distinct auditory stream in stage 2, and stream integration occurred in stage 3, at which time the streams gave rise to the perception of a single rich vocal tone. The application of pitch variation and vibrato (which may be described as amplitude modulation (AM) and frequency modulation (FM) synthesis) imparted parallel motion into the evolving frequency spectrum of the audio. This is but one grouping mechanism through which disparate spectral components may be perceptually fused with a sense of "belongingness" into a single auditory stream.

2.2.3. **Stream Segregation and Perceptual Grouping.**

The German word Gestalt means "structure" or "configuration" and this theory describes how the brain creates mental structures by interconnecting certain subsets of incoming sensory stimuli [65]. Auditory scene analysis, viewed through the lens of Gestalt psychology, becomes a task of establishing the relationship between figure (important spectral features) and ground (superfluous background noise). In the case of audified solar wind data, features that emerge within a narrow spectral bandwidth may be the result of significant wave-particle interactions, or they may be generated by instrumentally induced error—the auditory system must parse a set of complex and

[16] The father of Frequency Modulation (FM) synthesis [69].

potentially subtle spectral cues in order to make this distinction. This section will explore research conducted into the mechanisms that mediate this type of high-level perceptual organization.

In a 1950 paper published by Miller and Heise it was demonstrated that two groups of alternating tones played at varying frequencies resulted in either the perception of a single event moving up and down, or two separate notes, depending on the disparity in frequency between the two groups. It was noted that, "If frequency and time in the tonal pattern are replaced by vertical and horizontal spatial coordinates (e.g., as in the sound spectrographs of visible speech), then the principles of visual organization discovered by the Gestalt psychologists are applicable to the auditory situation." [71] This paper proposed a larger field of worthwhile research based on the investigation of auditory patterns; a baton that was picked up by Bregman in the 1970's. Bregman drew from the Gestalt principles of grouping in constructing a larger conceptual framework for how our auditory system creates higher-level organization from an array of complex environmental cues. These principles provide a set of rules whereby objects tend to "attract" one another based on **similarity, proximity, common fate, closure, good form, and good continuation** [65, 72-74]. While these principles were revealed through an examination of the mechanisms of visual perception, they readily translate to audition. **Figure 7** provides a visual reference for each of these principles (excluding good form).

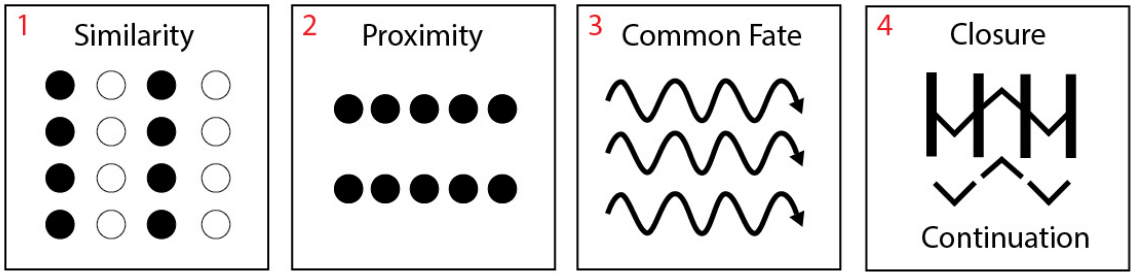


Figure 7. A graphical depiction representing the Gestalt principles of grouping.

The Gestalt principle of **similarity** provides a grouping mechanism between objects based on prominent attributes such as shape, size, and color. As displayed in **Figure 7**, the eye naturally organizes these sixteen circles into vertical rows—two black and two white. We may liken the principle of similarity to auditory timbre, as sounds with a similar timbre will be associated with a given group [65].

The second panel provides a series of circles with uniform color, in this instance physical **proximity** provides cues for grouping—the eye organizes the circles into two horizontal rows. Similarly, if we were to play back the circles as a sequence of musical notes (where the vertical spacing indicates separation in frequency, and horizontal space indicates time progressing from left to right) then the ear would tend to segregate the low and high tones into separate auditory streams.

The arrows in the third panel of **Figure 7** indicate parallel motion through space. A group of objects following this trajectory could be said to share a **common fate** (e.g., a large number of birds following this shared path will be visually grouped as a flock). From an auditory perspective we may consider these lines as representing a number of pure tones modulated up and down in frequency space through FM synthesis. The example of auditory stream integration provided in the previous section illustrated how such parallel motion gives rise to perceptual integration.

The principle of **closure** describes the perceptual mechanisms that tend to "fill in" missing information [65]. Bregman noted that the phenomenon of visual occlusion closely resembles auditory masking, where visual occlusion occurs due to the location of objects in physical space, while auditory masking occurs due to the relation of auditory objects in frequency space.[17] Observing the fourth panel of **Figure 7**—the visual modality infers the presence of a continuous line when the gaps between the three separate "v" shapes are occluded by black bars. We may also consider this figure from an auditory perspective as the spectrogram display of a non-continuous pure-tone swept up and down in frequency, occasionally occluded by broadband noise. In this example Bregman found that the auditory modality perceives a continuous pure-tone (despite the discontinuities) when the gaps are occluded by the noise mask [65]. This effect may also be described as the illusion of **continuation**. A set of audio examples was generated to illustrate these grouping mechanisms [75], and can be found online [76].

[17] The intensity of various auditory and visual stimuli may also be considered. A bright light will trigger an adaptive contraction in the eye, temporarily obscuring the perception of objects within a space. Similarly, a loud noise will obscure the auditory perception of subtle sounds within a space.

2.2.4. **Sequential Integration**

Sequential grouping also plays a significant role in the formation of streams. Miller and Heise investigated this effect with the most basic fundamental unit—a set of alternating pure tones—and revealed perceptual grouping as largely frequency dependent [71]. Bregman and Campbel extended this work in 1971, documenting that a single sequence of tones may appear to "break up" into multiple parallel sequences as the speed of playback increases [77]. Ongoing research has continued to revealed the significant impact of time-based effects in the role of perceptual grouping [65, 78]. **Figure 8** provides a visual depiction for several of these time-based mechanisms in action.

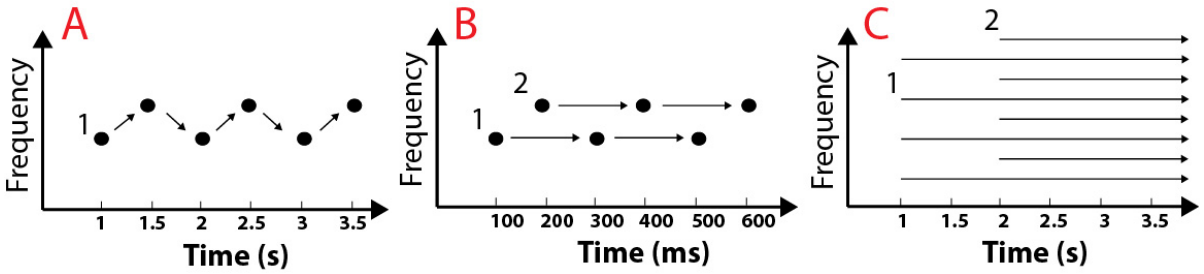


Figure 8. Three examples of pure-tone clusters exhibiting **A)** no stream segregation, **B)** segregation based on sequential integration, and **C)** segregation due to onset asynchrony (adapted from Bidet-Caulet 2009).

A sequence of alternating pure tones clustered within relatively close frequency proximity will result in a single auditory stream when played at slow speeds (**Figure 8, Panel A**). However, increasing the rate at which this sequence is played will significantly impact the perceptual grouping process, causing this single stream to segregate into two separate streams (**Panel B**). Bregman, who refers to this effect as the "illusion of multiple streams," found that in the case of more complex tone sequences the effect persisted even when the intensity of secondary "distractor" tones was reduced by 30 dB—this suggests that frequency-proximity takes precedence over loudness in the segregation of streams as the rate of temporal progression increases [66]. In light of this effect, a researcher exploring data through audification may decide to vary the sampling rate of a sound file to gain a sense for how the speed of playback impacts the perception of embedded features.

Stream segregation may also be considered in the case of harmonic timbres. Recent research demonstrates the perception of separate streams triggered by onset

asynchrony—when a group of harmonic pure tones is played and a second group is added several seconds later, the new tones will likely form a separate stream [78]. However, in the same situation, when this second set of tones is added with an onset asynchrony of less than 30 ms, the entire cluster is more likely to cohere into a single auditory stream.

2.3. Sensory Perception and Multimodal Interaction

The senses are traditionally assessed in isolation, from a modular standpoint. Researchers are only beginning to understand many of the complex mechanisms for the reception, encoding and decoding of stimuli. A large portion of the work presented in this dissertation investigates the abilities of the auditory and visual modalities when applied in tandem toward the identification of features embedded in audified data sets. In approaching the third primary research question (related to the potential benefits of auditory display) this work also directly investigates how the visual analysis capabilities of a heliospheric research scientist are augmented when auditory feedback is provided. Henceforth, the term *multimodal* shall exclusively apply to the co-application of audition and vision, and *cross-modal* will be applied to describe the effect of one sense on the other. This dissertation formally investigates the potential benefits of multimodal data analysis, as well as the potential influence of cross-modal cues in this analysis process. This section will review research that has been conducted to date toward the understanding of multisensory (audio-visual) interactions.

The process of sensory segregation begins when infants are presented an array of non-compartmentalized sensory information, and the task of experiential learning begins [79]. At this young age, sensory awareness could be described as exquisitely synesthetic, as it has been demonstrated that infants possess strong synesthetic associations that dissipate with age [80]. Infants as young as 21–31 days of age show an excitatory response to similar audio-visual cross-modal cues [81] and an attentional bias towards cross-modally congruent stimuli has been demonstrated as early as 3–4 months [82]. Cross-modal matching of brightness and loudness has also been demonstrated in children at the age of five [83]. At the most primitive level of synaptic relays, there is a direct correlation across all senses between the intensity of a perceived stimulus, and the resulting intensity of neuronal firing [84].

2.3.1. **Cross-Modal Effects on Spatialization**

The relative weight of spatiotemporal proximity in the creation of a coherent, unified sensory object cannot be downplayed. The *ventriloquist effect* describes the case in which our auditory perception is localized to the mouth of a dummy primarily because this motion provides such a strong visual cue. This is an instance of what is referred to as *visual capture*, in which our visual modality takes precedence in the presence of auditory cues that may be somewhat spatially incongruent [85]. Even if a loudspeaker is not placed directly in front of a projection screen, our sensory system will usually localize the voice of an actor to the moving image. It has been found that the cross-modal localization bias created in this instance will increase as a function of the disparity between the relative locations of the stimuli [86].

Such cross-modal correspondences are usually bidirectional [87], and research has shown that illusory visual motion can also be induced in the presence of temporally synchronized tones with alternating frequencies [88]. Auditory frequency also presents a strong cross-modal integration with visual height, as has been repeatedly demonstrated through research utilizing speeded classification tasks [89-91]. Specifically, the presence of cross-modally congruent stimuli (high auditory pitch & high visual location) reduces reaction time in comparison to cross-modally mismatched pairs (high auditory pitch & low visual location).

2.3.2. **Cross-Modal Effects on Perceptual Grouping**

It has been demonstrated that cross-modal information between the auditory and visual modalities can modulate perceptual grouping mechanisms. That is to say, our visual grouping schema may be altered by the presence of covarying auditory sequences, and again this effect is found to be bi-directional: a visual sequence alternating from low to high may affect our auditory perception of a tonal sequence with a somewhat similar low to high alternation [92].

The illusion created by Sekuler and Sekuler is a powerful demonstration of cross-modal spatiotemporal interactions at work [93]. In this example, two objects were presented that moved with crossing paths. It was found that sounds presented exactly at the point of coincidence increased the probability of perceiving the objects as “bouncing”

off of one another. An effect that was mitigated by the complete absence of sound, or in instances where sound was played *before* or *after* the moment of coincidence. These results are a clear demonstration of the ability of aural cues to transform our perception of visual motion.

2.3.3. **The Formation of Cross-Modal Associations**

The Bayesian approach to multimodal cue integration suggests a process of learned associations. From this standpoint, perceptual integration arises from iterative neuroplastic optimization based on probabilistic evaluations of environmental stimuli. A prime example is found in the cross-modal integration found in visual size and auditory pitch (where lower pitch is associated with larger object size). It has been demonstrated that this integration is absent in infants and is slowly acquired over time. It is suggested that this is a learned psychological association constructed through interaction with physical objects that demonstrate similar innate cross-modal behaviors [87]. A larger timpani-drum, for instance, will produce lower frequency tones when struck in comparison to a smaller drum. Ernst (2005) proposes that "the benefit of integrating sensory information comes from a reduction in variance of the final perceptual estimate" [94]. Given vast amounts of data from multiple sensory modalities, the human sensorium is remarkably adept at the process of reducing complex environmental stimuli into a discrete set of unified objects. In the case of auditory perception, an expanding body of research is beginning to uncover how this modality may contribute to the process of data analysis and knowledge extraction.

2.4. Sonification

The traditionally accepted definition of "sonification" (provided in the introduction) refers to the use of non-speech audio as a means of conveying information [1]. De Campo (2007) offers a more explicit definition for data sonification as "the rendering of (typically scientific) data into (typically non-speech) sound designed for human auditory perception" [95]. This definition is useful as it simultaneously embraces *human perception* while introducing *design* as a central component of auditory display. The most widely used sonification technique, known as **parameter mapping**, is a process through which data values are mapped directly onto parameters for sound

synthesis. A subset of this technique generates **auditory graphs** to map multiple data dimensions to sound synthesis parameters, and effective mappings often appeal directly to the intuition of the human operator [96]. **Auditory icons** are created from recorded sounds that are intended to mimic the physical analog of a digital action (for example, a user placing a file in the trash may hear the sound of crumpling paper). **Earcons** are generally short auditory cues, and are designed to alert a user as to a specific functionality. **Visual substitution** concerns itself with replacing, substituting, or enhancing visual stimulus. Through **model-based** sonification, data are transformed into a sounding instrument to be played by the listener [15]. This section will provide historical context for the field of sonification, and review the common techniques listed here.

2.4.1. **A Brief History**

The early investigation of auditory display was pioneered by Pollack (1954)—this research demonstrated that multimodal stimulation could increase the rate of information transfer to a human operator [10]. Loveless (1970) confirmed this increased efficiency and explored the benefits of presenting redundant audio-visual stimuli [12]. Through the implementation of a study that tasked participants with the detection of sinusoids embedded within noise, Fidell (1970) observed multimodal stimulation to provide increased sensitivity, facilitating superior task performance [11]. Yeung (1980) extended this investigation to human pattern matching abilities and demonstrated that known visual-analysis methods were often inferior to auditory analysis in the representation of multivariate data [13]. Several additional experiments utilizing multivariate data were conducted by Bly, who found that sound facilitated an increased understanding of multivariate data sets [14, 97].

Stephen Frysinger published a paper on *Applied Research in Auditory Data Representation* in 1990 while working for Bell Laboratories [98]. This work touched upon psychoacoustic and musical considerations in auditory data representation, reviewed the body of sonification literature published to date, and provided a framework for future experimentation in the field—calling for an investigation into the types of analysis tasks that may benefit most from the use of auditory display, and suggesting that

correlation analysis may be an ideal means for comparing signal detection capabilities across various display methods.[18] In the following year, Scaletti introduced the Kyma system, which provided a valuable platform for exploring the use of sonification as a means for extracting information from complex scientific data [99]. The first meeting of the International Community for Auditory Display in 1994 drew together a number of formerly disparate research initiatives, providing a much-needed structure through which sonification emerged as a formal discipline. The proceedings from the second ICAD conference held in 1994 are now considered a cornerstone text within the sonification community [100]. A report published by the National Science Foundation in 1999 surveyed the state of the field, citing a number of successful use cases, and calling for the "recognition of sonification as a valid research area" [1].

2.4.2. **Techniques**

Though the established boundaries and terminology are not entirely rigid, the following approaches are recognized within the sonification community [15].

2.4.2.1. **Auditory Icons and Earcons**

Metaphors for human computer interaction underwent significant development in the late 1980's. Gaver pioneered the development of **auditory icons** as "caricatures of naturally occurring sounds… used to provide information about sources of data" [101]. This research grew into the *SonicFinder*, a software platform that unified visual processes with auditory correlates through a delineated mapping process [102]. The efficacy of this display method was evaluated in a study in which participants were asked to monitor activities in a simulated working environment. Here, auditory feedback was found to have a significant impact on participants' interaction with the system—enabling participants to more effectively monitor the environment and "diagnose problems" [103]. This work demonstrated that sound could play an important role in human-computer interaction.

Earcons utilize short non-metaphorical, synthesized-sounds to provide state dependent auditory feedback to the user [104]. Both earcons and auditory icons are

[18] The first study presented in Chapter III applies correlation analysis as a means for determining the level of similarity between assessments made through auditory and visual analysis methods.

utilized in nearly all modern software interfaces [105, 106]. This type of representation, that is primarily non-quantitative in nature, is generally not considered among techniques for auditorily rendering scientific data sets—a process that may be referred to as "auditory data representation" [98].

2.4.2.2. **Auditory Graphs**

Mezrich (1984) introduced an interactive audio-visual platform for the real-time analysis of multivariate time series data sets—offering a tool for supplementing, rather than replacing, visual graphing techniques [107]. **Auditory graphs** were introduced shortly thereafter as a means for providing a direct sonic translation of two-dimensional graphs. This technique was offered as a tool for facilitating data analysis within the visually impaired community [108]. The effectiveness of auditory graphs in educational settings has been examined in a controlled environment where it was demonstrated that participants displayed a high degree of accuracy in matching auditory graphs with their visual counterparts [109]. Ongoing research initiatives are continuing to explore the potential of assistive visual substitution systems [110].

2.4.2.3. **Model-Based Sonification**

Model-based sonification is a particular form of interactive auditory data representation in which user actions "excite" a system that evolves based upon a mathematical model, subsequently producing an acoustic response [111]. This technique incorporates a data space, model space, and sound space—where the "setup" determines the correlation between the data and model, while the "linking variable" connects the model with the sound. This technique was introduced by Hermann in 1999 [112], and has been successfully applied in a number of cases [113-116]. One such case is the model-based sonification of Markov-chain Monte Carlo simulations [117], an approach largely based on particle trajectories and realized using several different operation modes (referred to as Methods). The researchers found that they were able to identify "complex dependencies evident in a chain," and the model facilitated navigation and feature identification in multi-dimensional datasets. This technique has strong applications for exploratory data analysis, and a proposed research agenda suggests the establishment of a broadly deployable "toolbox" [111].

2.4.3. **Parameter Mapping**

A commonly accepted definition for parameter mapping sonification is "the association of information with auditory parameters for the purpose of data display" [118].[19] This technique encapsulates the generation of auditory data representation that may be considered inherently "musical" or "non-musical." While both approaches are scientifically accurate by definition, non-experts may misconstrue the former as an interpretive process rather than a translational process. Musicality may be considered from a standpoint of psychoacoustic optimization. For example—mapping a continuous time series to a musical scale may require some level of quantization that can result in significant data reduction while increasing aesthetic quality and perceptibility. The mapping between data dimensions and parameters for sound synthesis is dictated by a transfer function, which may or may not be continuous and may be linear or non-linear. While a one-to-one correlation between data dimensions and sound synthesis parameters is common, many parameter mapping sonifications make use of redundant mappings (e.g., mapping a single data dimension to both the pitch and amplitude of a sinusoidal waveform), while many incorporate several data parameters in the synthesis of a single sound object. Grond and Berger (2011) review several additional considerations for mapping topologies [118].

Dubus (2013) provides a broad survey of parameter mapping techniques applied in the assessment of physical quantities, surveying 179 scientific publications spanning twenty years [119]. Of these projects, a subset was selected for extensive review. Parameter mapping sonification was found to be the most widely adopted strategy for sonifying datasets with multiple dimensions. In regard to research aims and objectives—the largest number of projects applied sonification for the purposes of data exploration. The second most common application was art and aesthetics, followed by accessibility, motion perception, and process monitoring (in order of decreasing prevalence). The team applied clustering to categorize physical data types into 5 categories: Kinematics (e.g., location, velocity), kinetics (e.g., force, temperature), matter (e.g., density, mass), time

[19] An argument could be made that all forms of sonification involve the mapping of information and sound, and hence, fall underneath the umbrella of parameter mapping. Terminology continues to evolve within the community, and there is still some active discussion regarding the definition of the term "sonification" and its proper use [114]. This meta-discussion will not be explored in this dissertation.

(e.g., elapsed time, frequency), and dimensions (e.g., volume, size). These physical quantities were mapped to auditory dimensions categorized as pitch, timbre, loudness, spatialization, and temporality (e.g., tempo, rhythm). Pitch was found to be "by far" the most commonly used auditory parameter, while timbre and tempo were used in a relatively small number of projects surveyed. The majority of these sonification projects were enacted within the Max/MSP programming language, followed by PureData and SuperCollider. While this database provides an excellent cross-section of parameter mapping research conducted to date, the team decided against the inclusion of audification, though the inclusion of audification was recommended as a future extension of this work.

The set categories established by Dubus (2013) were inclusively constructed to encapsulate an extremely wide range of sonification mappings. However, in reviewing related literature, it may be more helpful to adopt the classification of Walker and Nees (2011), who suggest categorizing sonification into groups of **1)** alarms, alerts and warnings, **2)** status, process, and monitoring messages, and **3)** Data exploration [118]. This review shall consider the first category as encapsulating non-quantitative methods (i.e., auditory icons and earcons), while the second and third categories encapsulate parameter mapping approaches in real-time and non-real-time respectively.

2.4.4. **Parameter mapping for Process Monitoring**

Vickers (2011) provides a classification for the monitoring of auditory information as a primary or secondary task. Process monitoring may be considered **direct** when the information display is intended to be the primary focus of the operator. **Peripheral** monitoring describes a situation in which the displayed information is intended to be non-focal. **Serendipitous-peripheral** monitoring refers to an environment in which the information display is not considered crucial to the task at hand, but merely supplementary [120].

The Geiger counter is one of the most well-known and widely referenced instances of auditory display. The first reference of its use was in 1917 when audification was applied to translate the electrical impulses caused by ionized gas in the tube of the counter into a series of clicks [121]. This device enables the user to monitor ambient

radiation levels while keeping the visual modality free (as navigation is of primary importance in radiation rich environments). Hence this device would fall under the peripheral monitoring category. Ambient soundscapes have also been explored as peripheral monitoring systems [122], and peripheral monitoring has also been demonstrated to be successfully applied in tandem with auditory icons [123].

The field of clinical medicine is one instance in which situational awareness is crucial, and several researchers have studied the efficacy of auditory displays in medical environments. An auditory environment was developed in which subjects were asked to monitor 8 simulated data parameters, including body temperature, heart rate, and blood pressure. The authors of this study reported most subjects preferred the auditory display after some practice, and the simultaneous perception of multiple data parameters proved to be advantageous in comparison to traditional visual displays [124]. Parameter mapping has also shown potential to support the ambient monitoring of respiratory information in a clinical setting [125]. Baier and Hermann have demonstrated a successful application for the real-time identification of seizures in EEG data [126-128]. Recently, auditory feedback has been demonstrated as an effective means for improving gait regularity in the rehabilitation of patients suffering from Parkinson's disease [129].

Sonification has broad applications in motor learning as a form of real-time bio-feedback, and it is widely reported that kinesthetic knowledge acquisition can be enhanced through augmented displays (e.g., displays that incorporate visual, auditory, and/or haptic feedback) [130]. The effectiveness of auditory display has been demonstrated in ski instruction [131] and karate training [132], and several studies have explored the benefits of parameter mapping sonification in speed skating for both general improvement [133] and movement rehabilitation [134]. Schaffert (2010) worked with a team of elite rowers, implementing a parameter mapping approach in which boat acceleration was mapped to the pitch of a continuous sinusoid [135]. The real-time feedback was found to be effective in synchronizing rowers, increasing the average velocity of the craft, and extending overall distance traveled. Dubus (2012) provided a quantitative evaluation of various methods for sonifying rowing movements and it was suggested that future efforts should strike a balance between functional and aesthetic properties in the resulting auditory display [136]. Ongoing research is investigating the

effectiveness of real-time auditory feedback in the optimization of swimmers' movements (such as the crawl stroke) through the sonification of hydrodynamic pressure [137, 138].

The monitoring of computer network activity has applications in cyber-security, where the transfer of information is often critical. Ballora (2011) explored this avenue as a potential means for distributing cognitive load across modalities [139]. Research in this area demonstrates that visual tools can be enhanced when the auditory modality is engaged [140]. In the language of scene analysis: when audition and vision work in tandem to analyze sensory stimuli, our attention mechanisms become more efficient at sorting relevant and irrelevant information.

2.4.5. **Parameter Mapping Sonification for Data Exploration**

This review will now shift toward a set of applications that more closely mirrors the audification research conducted in the body of this dissertation. This presentation is not intended as a comprehensive review of the field, rather, a number of successful use cases have been selected that exemplify the use of parameter mapping sonification in the investigation of scientific datasets.

Ballora (2004) considered heart rate fluctuations, as measured by an electrocardiogram (ECG), as a one-dimensional vector [141]. Preliminary research suggested that auditory display might be favorable to visual representation. This method was tested across 30 datasets displaying healthy (normal) heart rates, and 30 that displayed higher variability due to obstructive sleep apnea. A correct identification rate of 90% was achieved through the assistance of auditory analysis, and the authors subsequently suggested several clinical applications.

2.4.5.1. **Collaborative Parameter Mapping in the Sciences**

Alberto De Campo conducted a wide range of research collaborations toward the development of a flexible sonification software environment [142]. This interdisciplinary research included projects rooted in neurobiology, theoretical physics, sociology, and speech processing. In this work, the sonification specialist designed a number of specialized tools for sonification within the SuperCollider programming language. These software interfaces allowed for real-time interaction with parameter mapping and model-

based sonifications. Within theoretical physics, the sonification tool was determined to be a practical method for exploring the mass spectra of baryon atoms. Early documentation of this research noted that the complexity of the work had been underestimated, and scientists appeared to exhibit skepticism as to the feasibility of the approach. The work was extended into a doctoral dissertation, where it was noted that within the realm of physics, "the implemented sonification designs were not fully tested by domain experts in this quite specialized field" [143]. However, a number of user tests were conducted with experts in the iterative development of a tool for the assessment of EEG data (the test group of 4 participants was identified by the author as "small"). Qualitative data was collected in the form of survey results, which were used to inform the design of subsequent iterations. It was concluded that extensive training would be necessary for successful clinical applications.

This investigation did not produce a domain specific scientific outcome, though De Campo noted several important lessons learned while working closely with domain scientists. These include the suggestion to adopt domain-specific terminology whenever possible, keep interfaces as simple as possible, provide visible support for "what's going on," and to ensure users have adequate time to learn the software. It should be noted that many of these suggestions closely mirror the set of usability heuristics provided by Nielsen (1994) [144].

2.4.5.2. **Parameter Mapping of Space Physics Data**

In collaboration with researchers at Goddard Space Flight Center, Diaz-Merced (2008) applied the xSonify software platform to sonify data from the Radio Jove Receiver, which measures the presence of "plasma bubbles" in the ionosphere. Twelve ten-minute samples of 1 Hz data were gathered, and the power spectra were assessed through parameter mapping sonification. Specifically, the x-axis was mapped to time, and the y-axis was mapped to frequency space, such that the spectrum was played back with a steady tempo and amplitude, and a varying pitch. One instance of a connection between pitch and ionospheric recombination was noted, indicating the auditory detection of a

plasma bubble event. This technique was suggested for use in the visually impaired community [145].[20]

A pitch-mapping approach was applied in the evaluation of time series generated from a light curve of EX Hya, gathered by the Chandra X-ray Observatory [21]. In this instance, pitch was determined by phase, frequency, and time variations. Statistically significant regions in the data were identified through the extraction of harmonics. The power spectra of data from ACE, *Wind*, and GOES satellites were also sonified to analyze the impact of a major solar flare event, known within the heliophysics community as the "Halloween storms." It was noted that sonification "showed both expected and unexpected changes in the power spectra." (ibid p. 136).

This work was extended in a doctoral dissertation, in which Diaz-Merced conducted several workflow analyses and perceptual studies with research scientists at NASA GSFC [146]. Diaz-Merced focused on the ability of participants to recognize signals in simulated double-peaked "black hole" patterns that were visually obscured by Gaussian noise—these data sets generally contained hundreds of thousands of data points. This analysis task was chosen as it mirrored that of astronomers searching for indicators of black holes. Four analysis conditions were tested—visual only, audio only, audio-visual, and audio-visual combined with a red-line sweeping across the visual display. The study evaluated the signal-to-noise threshold at which participants correctly identified 75% of events. For the listening tasks, the sonification was generated with parameter mappings including waveform type (sine or square), stereo panning, amplitude fading, amplitude modulation, and frequency modulation, with the audio signal generated at a sampling rate of 22,050 Hz.

Repeated-measures analysis of variance (ANOVA) did not find significance for any condition in the first study, which did not include the use of the sweeping red-line. The second study increased the complexity of the embedded stimuli, and added the red line moving across in a linear fashion—ANOVA revealed that the audio condition led to better performance than the visual, and the audio-visual condition also outperformed

[20] This work was extended over the course of a doctoral dissertation and a potential feature of interest was identified within the data, though it was noted that this result has yet to be scientifically verified [143].

visual analysis. The third and final study supported the results of the second study, that auditory analysis may provide increased likelihood of signal detection.

Aside from the use of different sonification methods, the perceptual studies presented in this dissertation differ from the work of Diaz-Merced in several key areas:

1. Research reported in this dissertation investigates the application of *audification* methods while Diaz-Merced explored various *parameter mapping* approaches.
2. In the majority of analysis tasks, participants were provided with time series gathered from satellites as opposed to simulated data sets.
3. The high rate of information transfer through direct audification enabled the survey of longer time series.
4. Task completion time was measured in all instances.
5. Open-ended tasks provided information in regard to the types of assessments made across the two modalities.

2.4.6. **Tools for Sonification**

Currently, several tools for independent sonification research are distributed freely. *xSonify*, a cross-platform system developed by Candey et al., is provided by NASA GSFC as a sonification tool for space physics research [147]. *Sonification Sandbox* (written in Java) was developed by Walker et al. at the Georgia Institute of Technology as a toolkit for generating auditory graphs. This software is capable of translating a variety of datasets into MIDI files (a common computer-music file format for saving compositions as event sequences) [148]. The *Sonifyer* toolkit is currently available for the Macintosh platform and supports both sonification and visualization of data [149]. The *Sonipy* framework provides a set of modules for sonification within the Python programming language (an open-source platform) [150].

While many approaches for integrating audiovisual considerations into the interface design process have been proposed [95, 151-153], Stockman and Fraunberger noted, “The extensive design knowledge on perceptual mapping is not connected with high level interaction design.” Preliminary research has shown that multimodal cues have the potential to effectively shift the attention of operators, and potentially improve

performance in certain data analysis tasks [154]. However, investigations have largely been conducted on a case-by-case basis, and hence multimodal design knowledge has remained largely compartmentalized [151].

2.5. Applied Audification

Audification, as introduced in Chapter I, is the direct translation of data into audio samples. This section will track the emergence of early audification research at the turn of the 20th century, review a number of early studies that investigated the application of audification for the purposes of data analysis, and explore the history of audification in the space sciences.[21]

2.5.1. **Early Origins**

The advent of direct auditory monitoring as a means of data acquisition can be traced back to the invention of the stethoscope in 1816, an innovation that opened the doorway for auditory data analysis as a medical diagnostic tool [155]. Modern sound transmission technology came into being in 1876 with the advent of the telephone. The German physicist Emil Du Bois-Reymond published electrophysiological research in which he utilized the sound of his own voice to excite the muscles in a frog leg, at which time he reported "Evidently, the nerve seems to be more sensitive to some sounds than to others" [156, 157]. The advent of audification as a method for auditory display has been traced back to a series of papers published shortly thereafter in 1878, the first of which was published by Ludimar Hermann, in which audification was used in the exploration of electrical signals from muscles [158, 159]. Shortly thereafter, audification was officially proposed as a method for evaluating small electric currents [160]. In a paper published in the same year, Johannes Tarchanow documented the audification of both frog and human muscles. Tarchanow provided proof that muscle contractions could be audified, and employed a second telephone as a means for comparatively evaluating multiple audio signals [159, 161].[22] By the year 1900, technology for visualizing this type of electric signal was readily available, however, researchers continued to attest to the audification

[21] A set of audification methods is provided in Chapter VI, derived primarily from observations gathered during the case studies presented in Chapter V.
[22] Dombois (2008) and Volmar (2010) provide an excellent review of early audification research.

of EMG data as "virtually irreplaceable" [157, 162]. The advent of amplifier technology spurred advances in EMG audification techniques from Rudof Höber [163] and Ferdinand Scheminzky [164] in the early 1900's. A full review of early EMG audification practices is beyond the scope of this dissertation—the reader is referred to Pauletto and Hunt (2006) for a recent investigation into the potential benefits of EMG audification [165]. At present, many EMG and EKG technicians utilize auditory feedback as a means for diagnostic assessment, however, this practice is not well documented within the literature.

2.5.2. **Geo Seismology**

Audification research in the 1960s was primarily driven by auditory seismology. Early research conducted at Bell Laboratories explored the listener's ability to distinguish between earthquake sounds and sounds produced by bomb explosions. It was inferred that auditory data analysis could lead to successful classification, which proved to be true with a 90% success rate. It was also demonstrated that through the assessment of time-compressed recordings listeners were rapidly able to "separate significant events from the background, and to give a preliminary evaluation of their source" [166]. Further research similarly utilized earthquake and explosion sounds; trained observers in this study were able to successfully classify two-thirds of presented seismic signals [167, 168].

A foundational paper was published by Hayward in the proceedings of the 1992 International Conference on Auditory Display entitled *Listening To the Earth Sing* [16]. Here it was documented that domain scientists in the field of geo-seismology commonly sifted through archives of seismic data stored on FM tape recorders and listened back at high speeds to identify earthquakes. Hayward observed a striking disparity between the widespread use of this technique and the relative absence in the literature, noting that, "from discussions with experienced seismologists of the 1960s, this seems to have been quite common, although it is not well documented." It was noted that some scientists displayed the ability to classify seismic events more quickly through auditory analysis than through visual means [16]. In light of this fact, Hayward offered auditory display as a potential avenue for cost reduction when training new seismometer-technicians. More

recent research in the field of geo-seismology suggested that "the ear is able to challenge the epistemological powers of the eye" [17].

2.5.3. **Radio Astronomy, Whistlers, Tweeks, Chorus, Hiss, and Lion Roars**

Long before the proliferation of visualization technology, auditory analysis provided scientists with a method for the rapid evaluation of frequency spectra. In 1932, Karl Jansky observed static of unknown origin while listening to very low frequency (VLF) radio signals [169]; he was later able to deduce this static as having origins outside the solar system [170], heralding the advent of modern radio astronomy. Penzias and Wilson (1965) later reported the presence of a subtle noise measured by a 20-foot antenna in Holmdel New Jersey. These low-amplitude fluctuations were in fact produced by the cosmic microwave background radiation—an "echo" of the big bang [171].

Space research scientists investigate a wide array of spectral phenomena on a day-to-day basis, and a large number of these features have names with origins rooted in early auditory observations. The most widely recognized of these features are undoubtedly the electromagnetic phenomenon known as "**whistler**" mode waves first reported by Barkhausen in a 1919 paper where he wrote, "at certain times a very remarkable whistling note is heard in the telephone" [172, 173]. This audified "whistling note" was generated by radio emissions from lightning in Earth's magnetosphere; more specifically, the frequency-dispersion of these emissions as they propagated along the Earth's magnetic field lines, between the two hemispheres [23].

Whistler waves have a characteristic "descending" tone—as electromagnetic dispersion propagates high-frequency waves more quickly than their low-frequency counterparts, a train of high-frequency impulses will arrive first at the VLF receiver. The amplitude of whistlers is generally greatest around 5 kHz, though they have been known to sweep as high as 35 kHz and as low as 350 Hz. Falling in frequency over the course of approximately one second, these waves will generally fade at a lower limit of approximately 1 kHz. Whistlers will commonly cluster in groups, known as "echo trains," [23] in which case the rate of descent of each subsequent whistler will be less than the previous. Occasionally the radiation will echo in the ionosphere, creating a "faintly musical or chirping" sound known as a "**tweek**" [23]. The title *whistler mode* is

now applied to a family of electromagnetic waves, including ion cyclotron whistlers [174], and this phenomena has led to the identification of lightning on Jupiter [175], Saturn [176], and Venus [177, 178].

"**Chorus**" or "**dawn chorus**" is a type of electromagnetic radiation that occurs at very low frequencies (VLF) within the spectrum, usually consisting of a multitude of tones that tend to last 0.1 to 0.5 seconds and occur at a rate ranging from 1 per second to 1 per 10 seconds. This distinct type of VLF whistler-mode emission in Earth's magnetosphere occurs in a range of frequencies from 100 Hz to 5 kHz [24, 179]. Early researchers noted that chorus activity "will very often occur in bursts starting from a background of little or no chorus, rapidly building up in intensity and repetition rate, and then tapering down again to background noise. These bursts… sometimes seem to be triggered by whistlers. The triggering effect is often observed over periods of about an hour, during which time every whistler received is associated with a definite enhancement of chorus" [179]. When these modes sweep upward they are referred to as "risers," when they fall they are referred to as "falling tones," and instances of a decrease in frequency followed by an increase is referred to as a "hook" [24]. The term "chorus" is attributed to K.W. Tremellin, with the predominant theory that he likened the sound of the phenomenon to "the twittering of birds in their dawn chorus so characteristic of the English countryside" [180]. This hypothesis is bolstered by the fact that the strongest occurrence of this phenomena is observed at approximately 06:00 local time by mid-latitude stations [24].

Another type of feature originally classified through VLF radio observations, known as "**hiss**," is a powerful class of auroral kilometric radiation (AKR) and whistler-mode emission [25] consisting of trapped plasma waves that are unable to escape Earth's magnetosphere [181]. It is documented that auroral hiss is "generated along the auroral field lines by intense fluxes of electrons precipitating into the ionosphere with energies in the range from a few hundred eV (electronvolts) to several keV" [181]. The earliest recorded description of this type of phenomena was documented by J. M. Watts in boulder Colorado, who noted that "one type of sound sometimes heard on receivers in the audio-frequency part of the electromagnetic spectrum can best be described as a 'hiss'

due to its subjective effect on the ear" [182]. This nomenclature was quickly adopted by scientists in the categorization of this type of feature [183].

Smith (1969) observed regions of narrow-band magnetic noise occurring in sporadic bursts in Earth's magnetosheath, occurring at approximately 100 Hz in data produced by the search coil magnetometer instruments on several of the Orbiting Geophysical Observatories (OGOs). It was reported that "when the noise bursts are played into a speaker, they sound very similar to a lion's roar" [184]. This phenomenon is now commonly described as a **Lion Roar**, and additional research revealed that this class of transverse electromagnetic VLF wave propagates in the whistler mode with "right-handed" polarization [26, 185].

2.5.4. **Voyager 2 Plasma Wave Observations**

Scarf et al. documented a successful application of audification in the 1982 paper *Voyager 2 Plasma Wave Observations at Saturn*. The authors observed a number of impulsive noise bursts through visual analysis of wideband data gathered by the plasma wave receiver. Analysis of the waveform data did not provide a compelling explanation for the anomalous activity, and it was noted that "it was not possible to explain… in terms of external plasma wave phenomena." Subsequently the decision was made to audify the data—upon listening to the resulting audio the authors remarked, "The sound recording derived from this waveform frame provides a convincing way to identify the source of the intense turbulence. We summarize the audio analysis by stating that the sounds, which resemble a hailstorm, are those of impacts on the spacecraft" [22]. This provides an excellent example of a case in which audification led to the successful classification of an anomalous feature that confounded traditional analysis methods.

2.5.5. **Current Initiatives**

The last decade has seen a growth in the infrastructure for amateur radio astronomy—in 2014 NASA's radio Jove program reported the sale of 1800 DIY kits for constructing VLF receivers in more than 70 countries around the world [186]. This equipment enables observation of various astrophysical phenomena, including "radio noise storms" produced by the interaction between Jupiter and its moon, Io. It is reported that these events sound like "popcorn popping, or like a handful of pebbles thrown onto a

tin roof" [187]. This equipment has been used by amateur radio enthusiasts to identify radio emissions from solar phenomena, such as sunspots and solar flares [188], the latter of which are auditorily observed as "a rapid hissing noise followed by a gradual decrease back to the original audio level" [187]. In addition to purposes of Education and Public Outreach (EPO), the research team behind the program states that it is "capable of real citizen science" [186].

Audification in the field of space physics is applied more commonly than the literature would suggest. In many instances, audification is discussed only in the context of EPO—in the case of the STEREO mission's radio and plasma waves investigation, interactive web-based sonification was proposed as a primary EPO component [189]. It is unknown how many researchers within the community apply audification out of curiosity and how many have applied this technique for analysis purposes, though it is known that a number of audified satellite data examples can be found online [190, 191].

2.6. Conclusion

This chapter provided a broad history of sonification and audification techniques, with a specific focus on applications for data mining and analysis. It also presented a brief introduction to human audition, as the limits of this perceptual system play an important role in the evaluation of auditory data. The literature of Bregman's Auditory Scene Analysis and Gestalt theory may be referenced to draw parallels between the perceptual grouping mechanisms at work in the visual and auditory modalities, but this will provide limited insight as to how they perform in applied data analysis tasks. For this, it is helpful to emulate real-world scenarios that a research scientist might encounter, in a controlled study environment. The following chapter contains a series of studies that formally investigate the first research question: *how do auditory analysis capabilities compare with visual analysis methods in the evaluation of long time series datasets?*

CHAPTER III

Establishing a baseline for auditory analysis capabilities

This chapter contains three sections in which the field of heliophysics is used as a case study to extract quantitative information about the nature of observations made through visual and auditory assessment, and to quantitatively and qualitatively assess the value of auditory and multimodal analysis methods. The first section presents unpublished research, replicating the experimental design of Pauletto and Hunt (2005) in a comparative evaluation of auditory and visual analysis techniques. The second section is taken from R. L. Alexander, S. O'Modhrain, J. A. Gilbert, and T. H. Zurbuchen (2014), *Auditory and Visual Evaluation of Fixed-Frequency Events in Time-Varying Signals*, proceedings of the 20th international conference on auditory display, New York [192]. The final section presents an unpublished investigation of the grouping strategies applied by individuals in the completion of a structured feature identification task.

3.1. Comparative Evaluation of Auditory and Visual Data Analysis Techniques

Vision has long been the dominant perceptual modality within the sciences, and new interfaces that integrate audification may benefit from the use of multimodal interaction. This investigation recognizes the groundwork that has been established in the comparison of auditory and visual analysis methods, and seeks to both replicate and expand upon previous research. Twelve participants were each asked to review twelve data sets both visually and auditorily, and to rank each example based on a set of five heuristics. The visual and auditory responses were then assessed for correlation.

This study builds on the work of Pauletto and Hunt (2005). Their investigation uncovered a strong correlation between observations made through auditory and visual

analysis methods in the evaluation of complex time series data [19]. Here it was demonstrated that participants are able to make similar assessments through the use of auditory and visual analysis methods, an important first step in establishing "equal footing" for the use of audification in the evaluation of scientific data sets (as data analysis techniques have heretofore been dominated by vision). Their study incorporated a participant pool drawn primarily from students and faculty in a media and electronic engineering program, and consequently it is unknown how their findings will transfer to a group of participants that includes domain scientists. The study presented here incorporates two distinct participant groups: one with a computer-music background, and another comprised of heliospheric research scientists. An analysis of the data gathered from the entire participant pool produced results that were similar to those reported by Pauletto and Hunt, while analysis across the two groups revealed new information regarding the nature of visual and auditory analysis capabilities in a population of domain scientists. A strong correlation between assessments made by the two modalities was found in four of the five heuristics, and one instance of moderate correlation is discussed in terms of underlying perceptual processes.

3.1.1. **Experimental Method**

3.1.1.1. **Hypothesis**

It is hypothesized that there should be a strong correlation between the average responses across all participants in the visual and auditory data analysis tasks (this hypothesis mirrors the original hypothesis of Pauletto and Hunt). Such a correlation would indicate that similar qualitative judgments can be made when data are presented either visually or auditorily.

3.1.1.2. **Participants**

This study utilized a total of 12 participants: 6 computer music specialists and 6 domain scientists. Both groups had experience working with spectrogram displays, though the nature of the respective experience was quite different. The participant pool included 10 males and 2 females. All but one participant self-reported average to above-

average hearing. This single participant suffered from unilateral hearing loss, and was able to assess the full spectrum of the data stimuli monaurally. This group varied in age from 22 to 41, with an average age of 28.5. Musical training varied from none, to 7 years or more, with a predictable bias in musical training toward the group of computer-music specialists. Of the participants, 2 had received a high school diploma, 2 had received a bachelor's degree, 5 had received master's degrees, and 3 held doctoral degrees. Four of the participants considered themselves to be sound editing experts, and 4 considered themselves computer experts.

3.1.1.3. **Stimuli**

Both the visual and auditory tasks required participants to evaluate a series of twelve data sets, assess for the presence or absence of five data features, and provide a ranking for each on a scale from 1 (lowest) to 5 (highest). As this study utilized solar wind data sets, the five key analysis parameters were determined in consultation with a senior scientist in the SHRG. All of the following were considered pertinent analysis criteria in the evaluation of heliospheric data sets:

1. Presence of noise
2. Presence of repetitive elements
3. Presence of clearly discernible frequency components
4. Presence of data gaps
5. Overall signal power/loudness

The solar wind is extremely turbulent in nature, and hence most data sets can be considered "noisy" to some degree. However, the exact meaning of "noise" was further defined on a scale from isolated narrowband noise, to broadband white noise. Repetitive elements often occur as a function of the solar rotational period, as active solar regions can persist across multiple rotations and create similar traces within the data. At the micro-scale this type of feature will give rise to a clearly discernible repetitive element, while these same features will give rise to clearly defined periodicities (frequency components) over the course of several years. Data gaps can occur for a number of reasons, including instrument malfunction, scheduled repair, or the presence of a gap between subsequent missions (OMNI data-sets can be generated across multiple instruments).

Twelve data sets were used in this experiment: 7 generated from satellite data gathered from multiple instruments, 4 synthesized with digital-audio production software, and one derived from a recording of terrestrial wind.

The 7 satellite data sets spanned a wide range of instruments across a variety of missions, and all were confirmed by a domain scientist to be typical of those that a researcher might encounter in the field. One example contained the *t*-component of the magnetic field measurements in the RTN coordinate system, as measured by the magnetometer on the Advanced Composition Explorer (ACE) spacecraft [193]. The audified data contained a high-pitched frequency component that was likely produced by the rotational period of the spacecraft. Another example contained the *x*-component of the magnetic field as measured in the GSE coordinate system, which faces directly toward the sun. This example contained a set of rhythmic pulses due to sweeping magnetic field lines.

One example was generated by data from the Fast Imaging Plasma Spectrometer (FIPS) on the MESSENGER spacecraft [194]. These data samples were gathered as the instrument orbited Mercury, and several contain periodic pulses as the satellite passed through the planetary magnetosphere.

Two data sets were generated from OMNI solar wind plasma speed data gathered between 1963 and 1970. These data contained large gaps as there were time periods during which no instruments were operational. This data set was provided twice for the purpose of testing internal consistency.

One data example was provided from the Solar Wind Ion Composition Spectrometer (SWICS) [195] on the ACE spacecraft [193], and was chosen based on the presence of a strong underlying hum that was observed beneath the broad turbulent spectrum. This hum is generated by the approximately 27-day solar (synodic) rotational period. The final solar wind data set was selected from the MFI instrument on the *Wind* spacecraft; it provided another example of periodicity caused by rotating magnetic field lines.

All four artificially generated data sets were created in the Logic Pro production environment—compressed pink noise and fixed-frequency sinusoidal waveforms were used to emulate the turbulent spectrum produced by the rotating sun. These stimuli were

each designed to accentuate the presence of 1 or more of the analysis parameters, including clearly discernible frequency components, and data gaps. Toward this end, great care was taken to emulate the spectral characteristics of time series produced by various sun-observing spacecraft.

The final data example contained a recording of terrestrial wind that contained both subtle frequency components and broadband noise. Data gaps were manually inserted into this example. **Appendix 1** contains a complete list of data examples, and a brief description of their most prominent features.

All visual stimuli were rendered in the iZotope Rx software environment, and consistent settings were used to provide uniformity across all examples. Spectrograms were presented with a logarithmic scaling on the x-axis, and a chromatic color mapping from black, representing the absence of spectral energy, to white, representing the full presence of spectral energy. This visualization method was reviewed by a member of the SHRG and deemed appropriate for the spectral representation of solar wind data. The same 12 stimuli were used in both the visual and auditory tasks.

3.1.1.4. **Experimental Procedure**

The experiment was conducted on a 15-inch MacBook Pro running the Mac OS X (10.6.7) operating system. The listening task was completed with Sony 7509-HD professional dynamic stereo headphones. The pre-test, analysis tasks, and post-test were encapsulated within a single standalone application constructed in the Max/MSP computer-music programming environment (version 6.05). All responses were recorded using the "coll" object and saved as data files in *.txt* format. A time-stamp for individual responses was recorded, along with total completion time for each task. Before beginning the experiment, participants were prompted to provide their first name, middle initial and last name; unique 3-letter file names were created from the initials.

All participants completed the entire test in one sitting. Testing was conducted at various locations at the University of Michigan that were quiet and free from distractions. The experiment was administered with an interface constructed in the Max/MSP computer-music programming environment. Participants were provided with headphones and given a brief verbal overview of the task. After completing a short pre-test

questionnaire, a randomization procedure implemented within the interface determined the ordering of the visual and auditory tasks.

Prior to beginning each task, a training module was provided to explain the significance of each data feature and introduce the participants to the experimental interface. One training module guided participants through the process of listening to auditory data, and the other provided assistance in conducting a visual analysis of a spectrogram (See **Figure 9** on the following page). The data file used for the training session contained at least one example of each feature. After each training module was completed, the analysis portion of the study began.

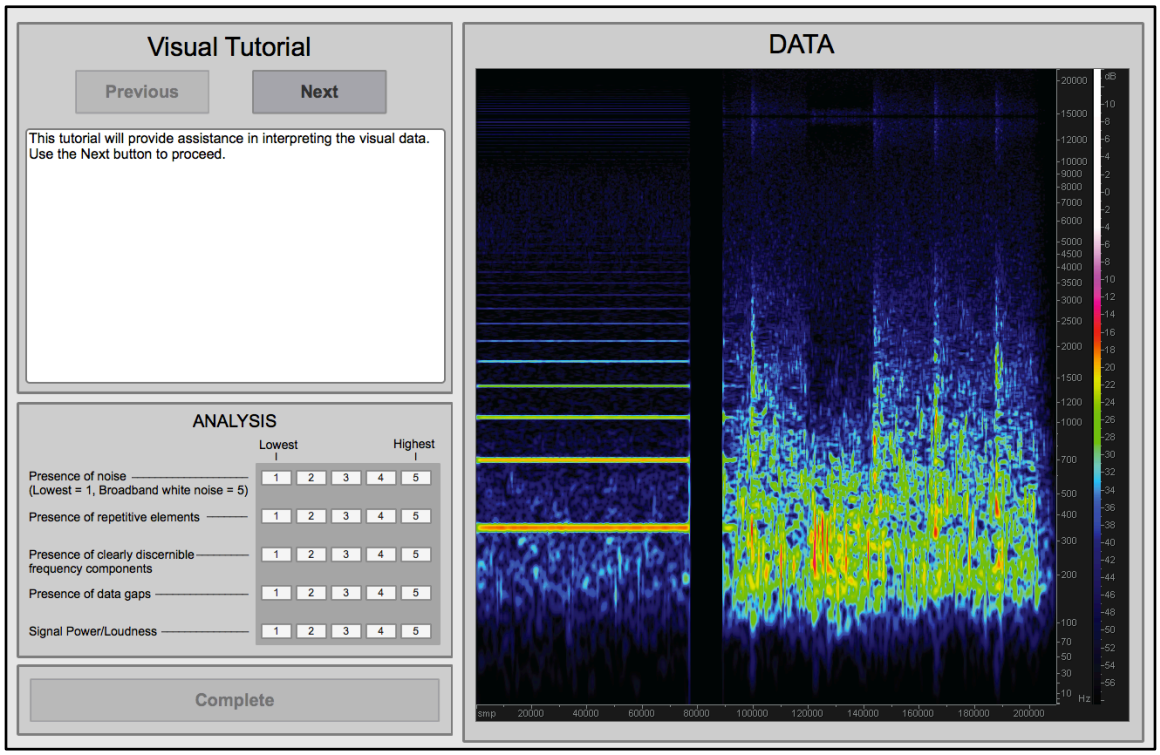


Figure 9. The visual tutorial provided an example spectrogram.

The stimuli were arranged on-screen in a 4x3 grid of thumbnails (**Figure 10**), and pressing a button labeled "open test" revealed a new screen in which the stimuli were reviewed and ranked. Participants were instructed to review all data examples before beginning their evaluation, and in the auditory task they were not able to change the volume setting once all stimuli had been reviewed. This prevented a potential perceptual bias that could be introduced through a global volume change partway through the

experiment. The order in which audio and visual stimuli were presented on screen was randomized prior to each examination. Participants were able to freely move back and forth between stimuli in any order of their choosing, and upon submitting responses for a given data example participants were no longer able to review or modify the submitted rankings.

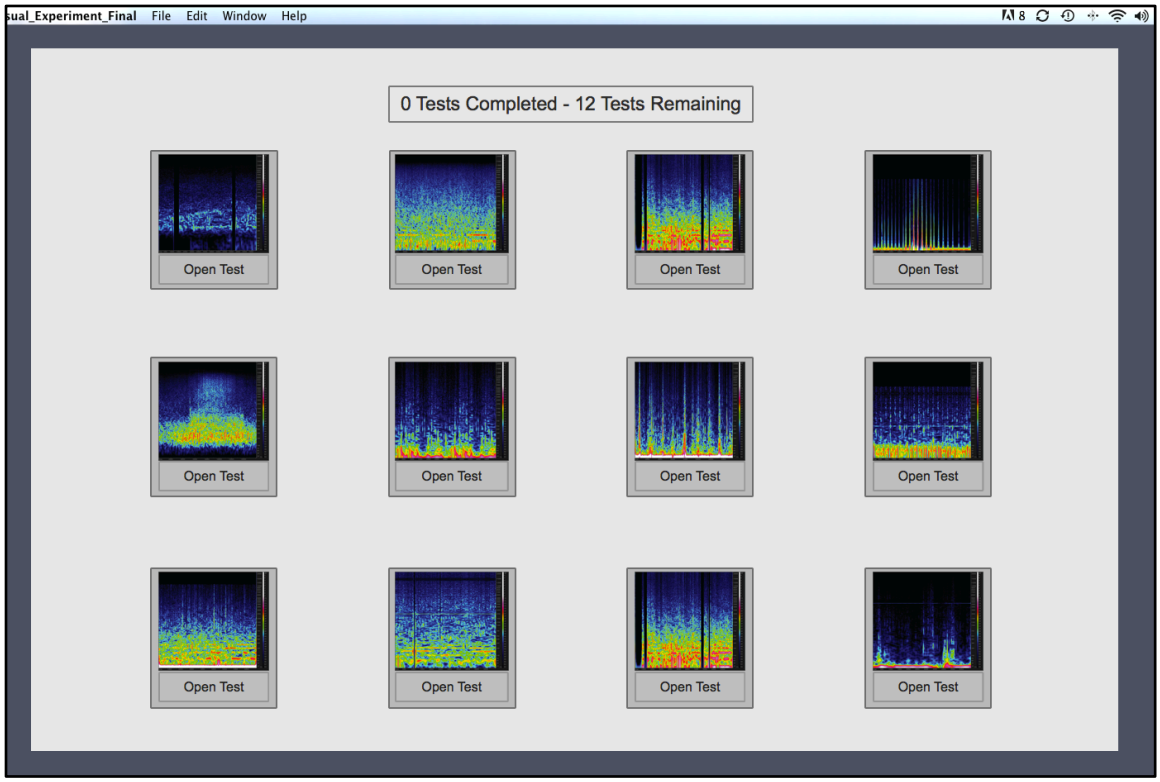


Figure 10. Data examples for the auditory and visual tasks were arranged on screen in a 4x3 grid.

For the auditory portion of the task, participants were provided with several options for starting and stopping playback. An on-screen play-bar could be used to start the sample from any specific location, and the space bar could be used to start and stop playback. Additionally, a looping option allowed participants to listen to each audio example repeatedly. It was noted in the tutorial that looping essentially causes the entire stimulus to become a repetitive element, and to be aware of this potential bias during the ranking process. The interface for the auditory task is displayed in **Figure 11**.

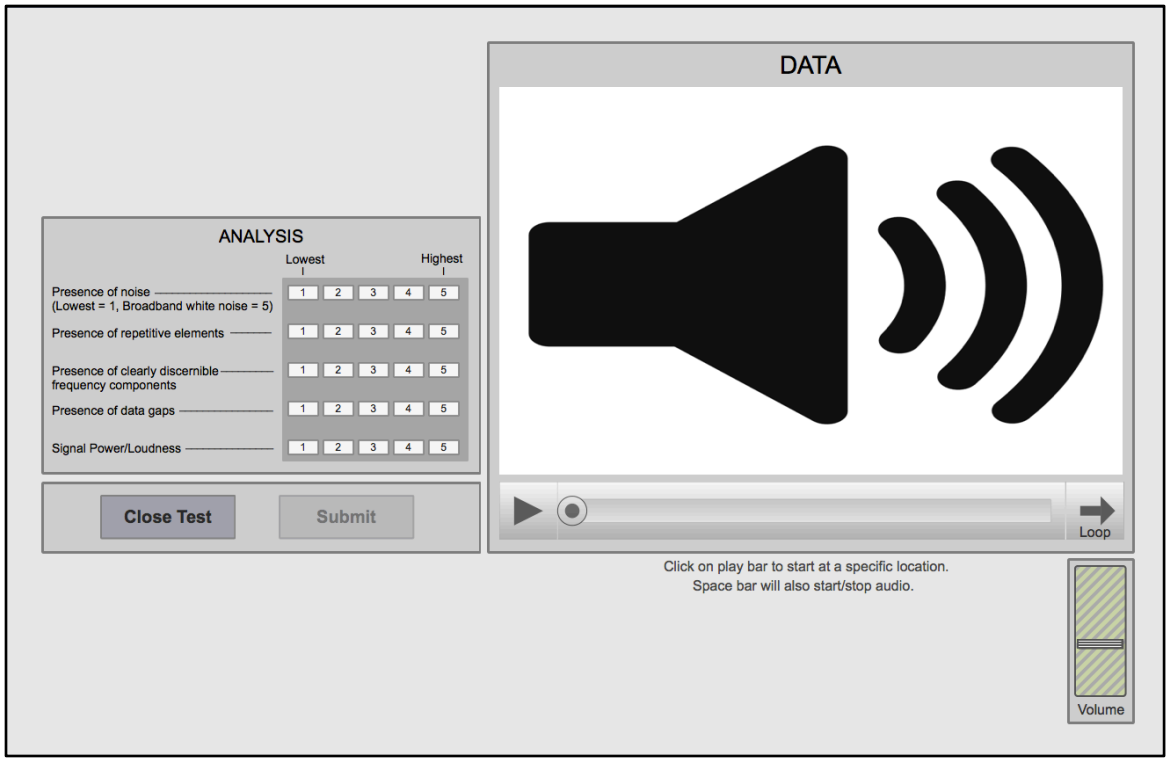


Figure 11. The interface for the listening task. The right side contains controls for audio playback and volume adjustment; the left side contains a form for recording participant responses during the ranking task.

A post-test questionnaire was specifically designed to determine participants' familiarity with sonification, experience working with spectrogram displays, level of comfort with computers, and experience with data analysis, mathematical modeling, and scientific research. This information was gathered in order to assess for a potential correlation between individual backgrounds, and performance on the analysis task. Participants were also asked if they noticed any duplicate audio and/or visual stimuli, and if yes, to write how many. While the duplicate stimuli could be used as a measure of internal consistency, a participant's awareness (or lack thereof) of these duplicate stimuli could also yield potentially valuable insight as to their cognitive state during the examination. Finally, a space was provided for additional feedback in free-response form.

3.1.2. **Results**

The average completion time for the auditory task was 6.9 ± 1.9 min, while the average completion time for the visual task was 6.2 ± 1.5 min. Half of the participants completed the auditory task more quickly, and the other half were faster in the visual

task. A matched two-tailed *t*-test did not reveal any statistical significance in the difference between completion times ($p = 0.15$).

Analysis of the task data was conducted in a manner that mirrors the methodology of Pauletto and Hunt, such that the results of the two studies can be directly compared. First, mean values were calculated across all participant responses for each data example (in order to determine the average response for each of the five features). One matrix was created from the visual responses, and another from the auditory. These two matrices were utilized to create *x*-*y* pairs in which *x* values corresponded to averaged scores for the auditory task, and *y* values corresponded to averaged scores for the visual task. A scatter plot was then generated for each of the five data features. As noted by Pauletto and Hunt, if participants had provided identical responses for both tasks, the resulting plot would contain a perfectly straight line (with a slope of 1). These scatter plots contain both a line indicating perfect correlation, as well as a regression line fitted to the data.

It is possible to determine the interrelationship between the two modalities by calculating Pearson's product-moment correlation coefficient [13]. This equation (1) is used to calculate linear dependencies between two variables. A resulting correlation factor of 1 indicates a perfect correlation between the two variables, a factor close to zero indicates very little or no correlation, while a factor of -1 indicates a perfect anti-correlation.

Correlation factor

$$r = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sqrt{\sum(x-\bar{x})^2 \sum(y-\bar{y})^2}} \qquad (1)$$

In all instances, correlation strength is reported as weak, moderate, strong, or perfect. These measures are based on the criteria established by Dancy and Reidy (2004), as shown in Table 3.

Table 3. Correlation strength, as calculated through Pearson's product moment correlation coefficient and defined by Dancy and Reidy (2004).

No-association	Weak	Moderate	Strong	Perfect
$r = 0.0$	$r < 0.39$	$0.40 \leq r \leq 0.69$	$0.7 \leq r \leq 0.9$	$r = 1.0$

3.1.2.1. **Presence of Noise**

The following scatter plot contains responses for the presence of noise averaged across all 12 participants for each example. Averaged auditory scores are plotted as a function of averaged visual scores. The solid black line indicates perfect correlation, and the dashed-line is a linear regression fitted to the data.

A strong correlation was found between visual and auditory analysis in scoring for the presence of noise (r = 0.80) (see **Figure 12**). This correlation is lower than that found in the scoring of data gaps, much higher than that found for power/loudness, and slightly higher than the correlation in the frequency component assessment task. The correlation is approximately equal to that of the scoring for the presence of repetitive elements. A high standard deviation was found across individual participant responses for each example (s = 0.29). Averaged scores were most similar between the two modalities for Examples 3, 5 and 6, and dissimilar for Example 2 (see **Figure 13**).

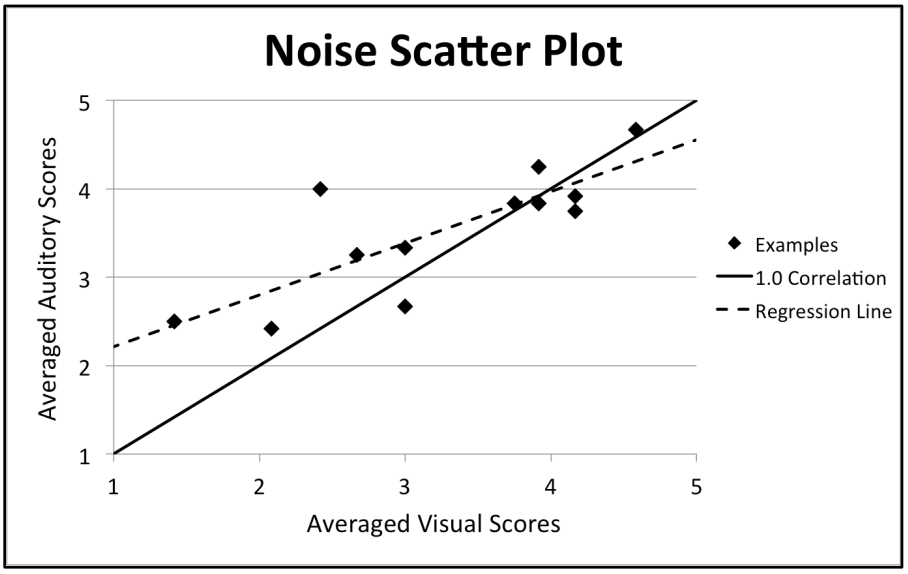


Figure 12. Scatter plot for presence of noise (averaged scores for each data example).

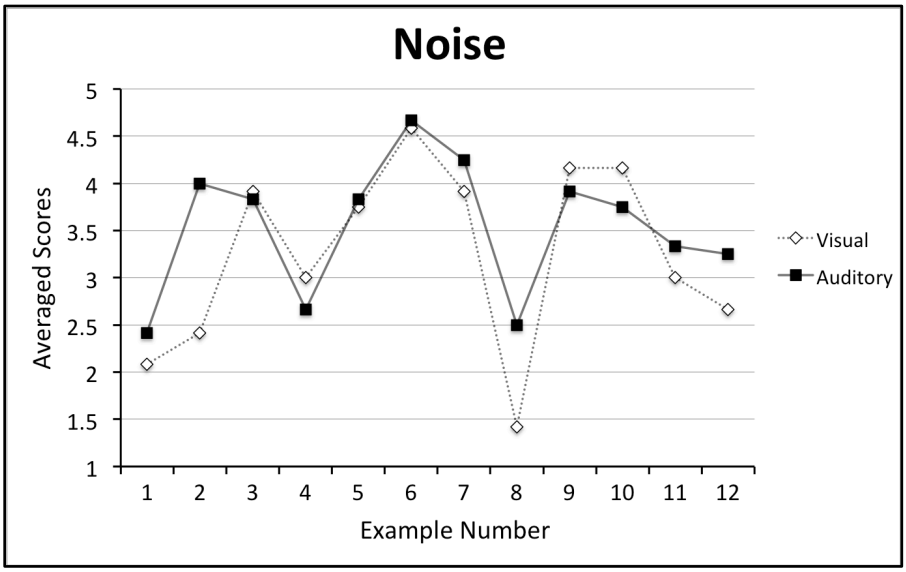


Figure 13. Presence of noise (averaged scores for each data example).

3.1.2.2. **Presence of repetitive elements**

A strong correlation was found between visual and auditory analysis in scoring for the presence of repetitive elements (r = 0.80) (see **Figure 14**). This correlation is lower than that found in the scoring of data gaps, much higher than that found for power/loudness, and slightly higher than the correlation in the frequency component assessment task. The correlation is approximately equal to that of the scoring for the presence of noise. A moderately large standard deviation was found across individual participant responses for each example (s = 0.22). Averaged scores were most similar between the two modalities for Example 8, and dissimilar for Example 5 (see **Figure 15**).

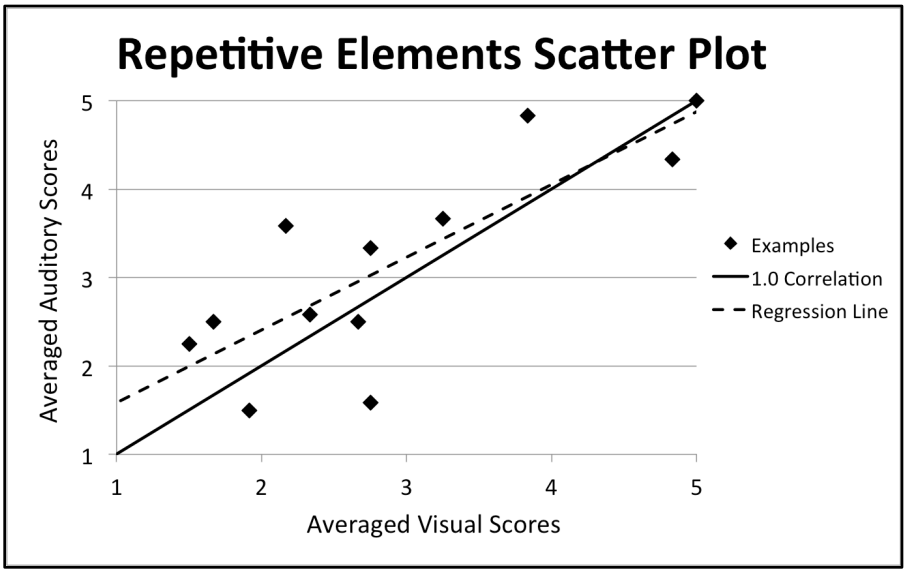


Figure 14. Scatter plot for repetitive elements (averaged scores for each data example).

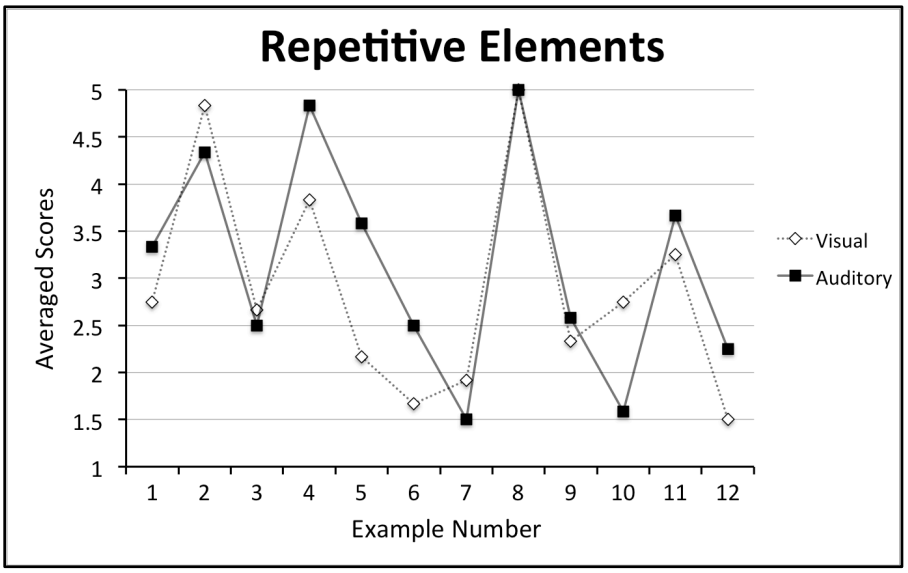


Figure 15. Presence of repetitive elements (averaged scores for each data example).

3.1.2.3. **Presence of clearly discernible frequency components**

A strong correlation was found between visual and auditory analysis in scoring for the presence of clearly discernible frequency components (r = 0.77) (see **Figure 16**). This correlation is lower than that found in the scoring of all other data features except the overall signal power/loudness. The standard deviation across individual participant

responses for each example was small (s = 0.12). Averaged scores were most similar between the two modalities for Examples 2 and 9, and dissimilar for Example 1 (see **Figure 17**).

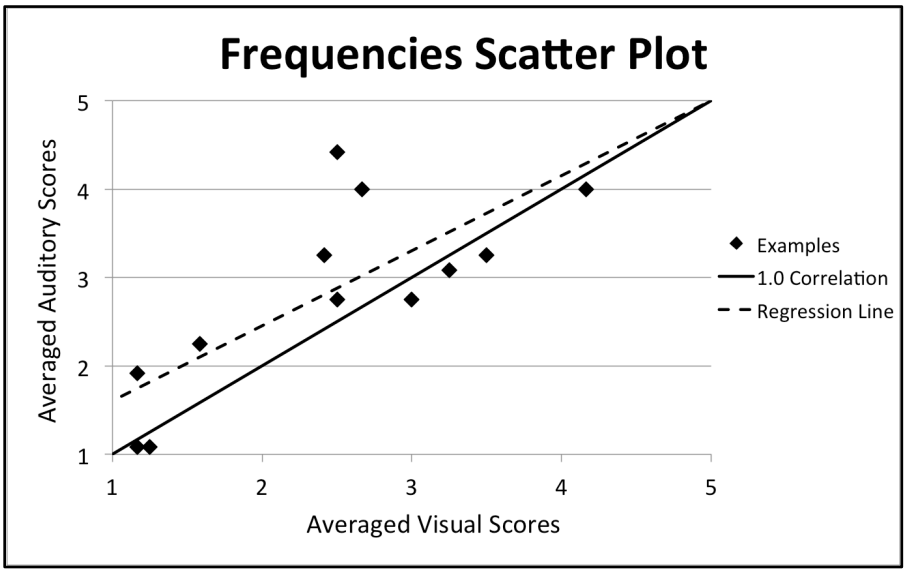


Figure 16. Scatter plot for frequencies (averaged scores for each data example).

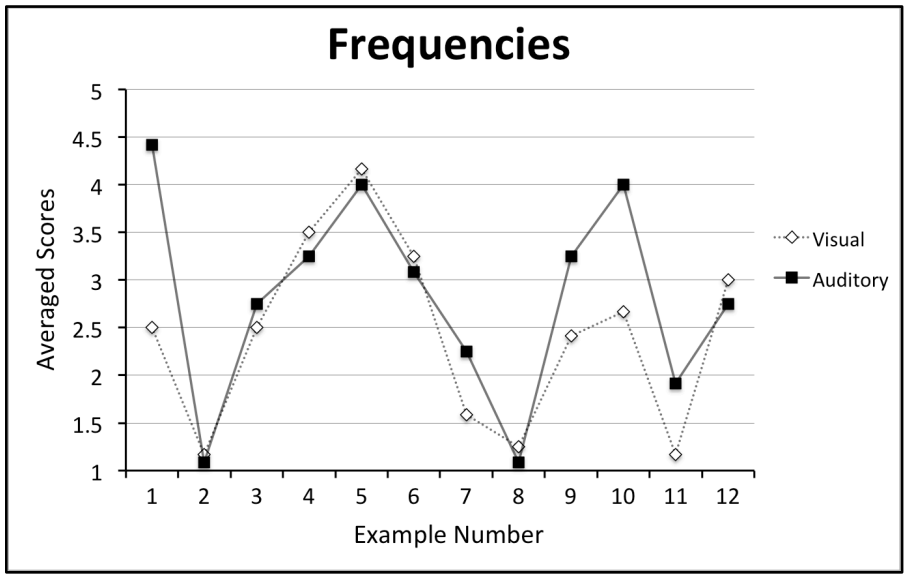


Figure 17. Presence of frequencies (averaged scores for each data example).

3.1.2.4. **Presence of data gaps**

A strong correlation was found between visual and auditory analysis in scoring for the presence of data gaps (r = 0.96) (see **Figure 18**). This correlation is higher than that found in the scoring of all other data features. The standard deviation across individual participant responses for each example was small (s = 0.14). Averaged scores were identical between the two modalities for Examples 6 and 7, and most dissimilar for Example 4 (see **Figure 19**).

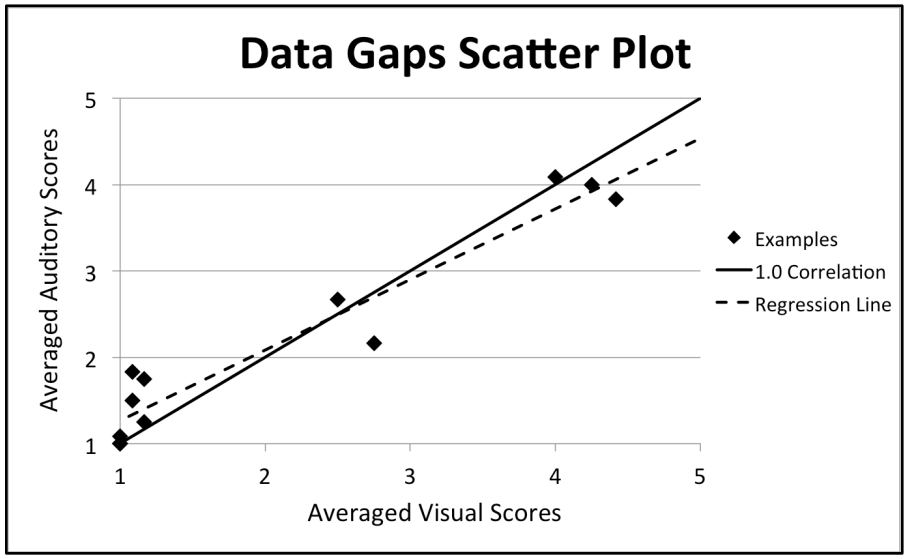


Figure 18. Scatter plot for data gaps (averaged scores for each data example).

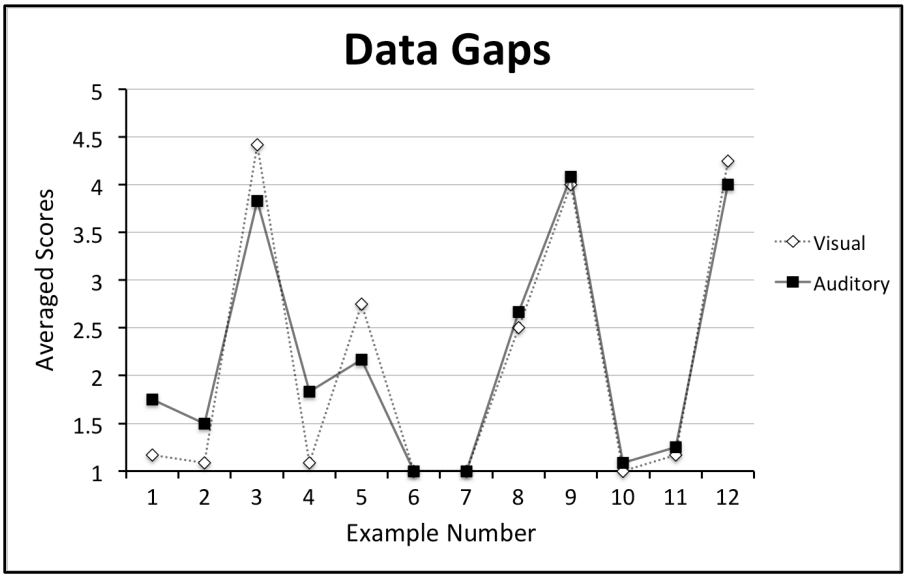


Figure 19. Presence of data gaps (averaged scores for each data example).

3.1.2.5. **Overall signal power/loudness**

A moderate correlation was found between visual and auditory analysis in scoring the overall signal power/loudness (r = 0.56) (see **Figure 20**). This correlation is lower than that found in the scoring of all other data features. The standard deviation across individual participant responses for each example was relatively large (s = 0.31). Averaged scores were identical between the two modalities for Example 7, and most dissimilar for Example 8 (see **Figure 21**).

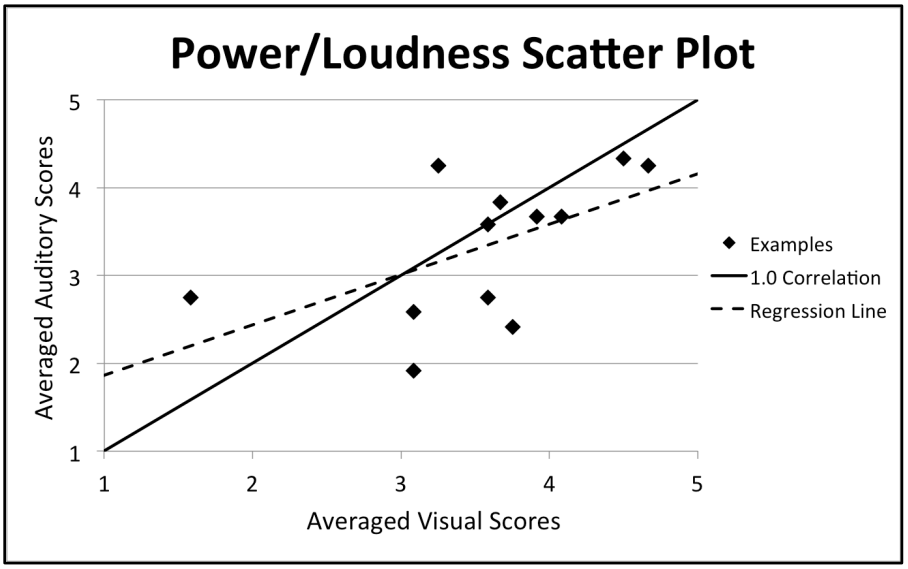


Figure 20. Scatter plot for power/loudness (averaged scores each example).

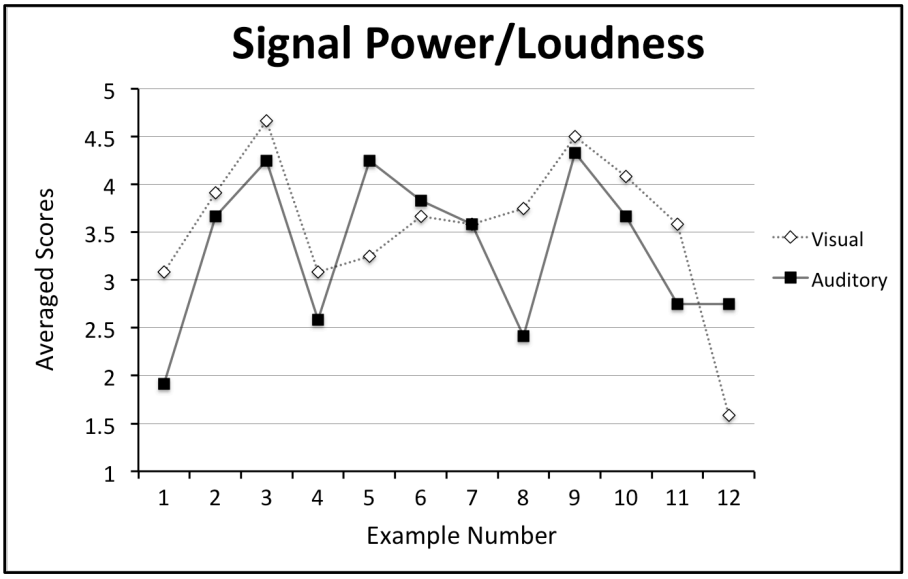


Figure 21. Overall signal power/loudness (averaged scores each example).

3.1.3. **Discussion**

The correlations found while mirroring the analysis methodology of Pauletto and Hunt (summarized in **Table 4**) were close to those found in the original study, with the exception of scores for data gaps (originally presented as "discontinuities") and the overall signal power/loudness (originally presented as "signal power"). The outcome of this study supports their original finding that the correlation between assessments made through visual and auditory means is generally strong. With respect to the profile of participants as obtained through the pre-test questionnaire, demographic information did not have a statistically significant impact on the degree of the correlation between rankings across the two modalities.

Table 4. Correlation across the two modalities, presented in order from strongest to weakest.

Presence of data gaps	$r = 0.96$
Presence of noise	$r = 0.80$
Presence of repetitive elements	$r = 0.80$
Presences of frequency components	$r = 0.77$
Overall signal power/loudness	$r = 0.56$

3.1.3.1. **Analysis of overall signal power/loudness**

The lowest correlation was found for the scoring of overall signal power/loudness. The root of this may be traced back to the physiology of the human auditory system where the frequency spectrum is parsed in a method akin to a Fast Fourier Transform with high resolution in both the temporal and frequency domains [196].[23] Critical band theory states that these cilia can be excited by stimuli within a certain frequency bandwidth, and that saturation of this bandwidth will result in a skewed perception of overall loudness. In this way, the auditory system may perceive broadband and narrowband noise to be at different overall power/loudness levels even if the overall power/loudness of the two signals is identical [197].

[23] See Chapter II for a review of the human auditory system.

Examples 1 and 5 presented a large variance in average scoring between the visual and auditory modalities in overall loudness/power. These stimuli both contained sinusoidal elements that fell within the most sensitive frequency range of the auditory system (roughly 500 Hz – 4 kHz). It is possible that these frequency components had a strong effect on the perception of overall loudness, and hence gave rise to a large variance in scoring.[24]

3.1.3.2. **Analysis of the presence of frequency components**

Examples 1 and 10 showed the lowest similarity in average scores between the auditory and visual assessment of discernible frequency components. This contributed to a lower correlation score than that of all other elements except for the evaluation of overall signal power/loudness. It should be noted that examples 1 and 10 both contain frequency components embedded within broadband noise, and in both instances the presence of frequencies was rated higher in the auditory modality than in the visual. This reinforces the notion that the auditory modality may pick up on spectral components that may be overlooked through visual analysis, as suggested by Hayward (1994). In this sense, the moderate correlation score is indicative of an area in which audification may offer strong support when applied in tandem with data visualization methods.

3.1.3.3. **Characterizing the two groups**

This study utilized a mix of computer music specialists and domain scientists. Both groups had experience working with spectrogram displays, though the nature of the respective interactions was quite different. The axes are of extreme importance within the heliospheric sciences, as the rate of temporal evolution and exact timing of a physical phenomenon have important implications. Similarly, the scaling of the *y*-axis (logarithmic or linear scaling) is a decision that is made with considerations as to the nature of the data at hand, and utilization of vertical space is often optimized for viewing phenomena of particular interest. Within the field of computer music it is generally

[24] It is known that the perception of energy within different spectral bands will vary depending on the overall loudness of the stimuli (see Chapter II). It is possible that an interface could integrate volume-dependent equalization that inverts the curves found by Fletcher and Munson. This could potentially lead to more accurate representations of loudness with stimuli presented at varying levels of amplitude.

assumed that spectral displays will render frequency content from 20 Hz – 20 kHz, and the y-axis will span a temporal duration ranging from a few seconds to several minutes.

Members of the SHRG report that spectrogram displays are often utilized in the visualization of highly turbulent data, where the analysis of frequency content is of high importance. This group generally does not utilize interfaces in which data processing techniques are applied directly to a spectral display. Processing is applied to the original data matrix, which is then rendered through mathematical operations such as FFT and Wavelet transformation, and the results of these processes are then inspected visually.

Computer music specialists typically interact with spectrogram displays within digital editing software that enables direct editing and processing of the audio waveform. Furthermore, these displays allow the user to move very fluidly forward and backward in time, and to zoom in and out freely. Within the computer music field, spectral visualization often comes secondary to audition, as the ear is the primary analysis tool within the domain of audio production and mastering. Nonetheless, spectrogram displays are extremely helpful in the process of locating and removing potential artifacts such as hum and clicks.

3.1.3.4. **Analysis Across the Two Groups**

In the primary analysis, correlation scores were calculated for each participant across all data examples. It should be noted that these correlation values are much lower than those found through the averaging of all participant responses for a particular data example. This is to be expected, as trends derived from average group responses will generally show a smaller variance than scores for individual participants. The following section explores the differences in averaged participant correlation scores across the two groups: one group containing six computer music specialists, and another group containing six domain scientists.

Analysis of the individual participant scores indicates a moderate correlation between auditory and visual analysis tasks in both the computer music specialists (r = .59) and the domain scientists (r = .58). In all instances, correlation scores will be provided first for the computer music group, and second for the domain scientists. Both groups achieved the highest level of correlation in the analysis of data gaps (r = 0.66 and

$r = 0.63$), while both groups achieved the lowest correlation in the evaluation of overall signal power/loudness ($r = 0.16$ and $r = 0.34$).

The computer music specialists showed a higher average correlation in the evaluation for the presence of repetitive elements between the auditory and visual modalities ($r = 0.58$ and $r = 0.42$). This group would have a large amount of formal training in the evaluation of musical rhythms, and consequently it is expected that this score would be higher. The domain scientists showed a higher average correlation in the evaluation for the presence of noise ($r = 0.35$ and $r = 0.51$). Space research scientists are familiar with the process of evaluating inherently turbulent and chaotic data, and hence it is possible that their sensitivity to various *types* of noise may be higher.

3.1.3.5. **Evaluating Internal Consistency**

No participants were entirely consistent across the repeated OMNI data example, and half of the participants reported in the post-test that they neither saw nor heard any repeated data examples. Three participants had a correlation of 1 across the repeated stimuli in the visual analysis task (there were multiple instances of a correlation of $r = 0.975$), while none achieved a correlation of 1 through auditory analysis. All participants who displayed perfect internal consistency in the visual task completed the repeated stimuli consecutively, while many participants with lower correlation scores evaluated other stimuli before returning to the second of the repeated stimuli (the order in which stimuli were assessed was controlled by the participants, as described in the experimental procedure). This lack of complete internal consistency likely reflects the evolution of individual scoring standards over the duration of the study.

3.1.3.6. **Evaluating subjective feedback**

All but one participant reported that they believed listening to data could provide valuable insight in the data analysis process. After completing the study, one domain scientist noted, “It was easier to pick things out with the ear than I originally thought it was going to be, especially the repetitive elements and the frequencies.” This sentiment was echoed by several scientists, who reported that they were surprised by the ease with which they were able to detect repetitive elements and frequency components in the audification.

3.1.4. **Conclusion**

This section presented a study that mirrored the experimental design of Pauletto and Hunt (2005) while incorporating a group of domain scientist into the participant pool. Participants were visually and auditorily presented with a variety of audified data examples, which were primarily gathered from various sun-observing spacecraft. Average scores for all participants across each data example displayed a strong correlation between auditory and visual assessments. Additional analysis was conducted between the correlations returned by computer music specialists and those returned by domain scientists. It was found that the scores were similar, though the computer music specialists displayed higher average correlation in the evaluation of repeated features, while the domain scientists displayed higher average correlation in evaluation for the presence of noise. Potential rationale accounting for these differences was discussed. This study found several areas of research that warrant further investigation, including a comparatively low correlation between the visual and auditory modalities in the analysis of overall signal power/loudness.

3.2. Auditory and Visual Evaluation of Fixed-Frequency Events in Time-Varying Signals

This study directly compares the auditory and visual analysis capabilities of participants in a structured data analysis task. This task involved the identification of transient fixed-frequency sinusoid events that were embedded within white noise and noise derived from solar wind time series. It was hypothesized that participants would be able to identify the number of embedded events more quickly and accurately through auditory data analysis than through visual analysis. While visual analysis outperformed auditory analysis overall, additional investigation revealed that auditory analysis outperformed vision in instances where these events were embedded in solar wind data. This task—involving the detection of transient periodic activity occurring within background turbulence—closely mirrors a type of spectral analysis conducted by heliospheric scientists. Additionally, several data examples contained embedded events that were correctly identified through audition while being consistently overlooked

through visual inspection. The largest disparity between visual and auditory performance was found in the analysis of white noise spectra that contained no embedded events. In these instances, auditory analysis regularly resulted in the identification of events when none were present; a potential reasoning for these false positives is discussed. The results of this study suggest that the analysis capabilities of each modality may vary based largely on the complexity of the masking stimuli that are present.

The goal of this study is to directly compare the auditory and visual analysis capabilities of participants in a structured data analysis task. The participants consisted of two groups at the University of Michigan who have experience working with spectral displays: heliospheric researchers and computer-music specialists. Transient sinusoidal waveforms were embedded in time-varying signals that contained background noise, and the task of the participant was to identify how many of these transient events occurred within each example. This is similar to a type of spectral analysis task found in the heliospheric sciences. It is hypothesized that participants will be able to identify the number of time-varying fixed-frequency sinusoid events more quickly and effectively through auditory display than through visual analysis. This section will provide a psychoacoustic context for this study before presenting the experimental design and significant findings. The results will be discussed, and finally, various avenues for future investigation will be proposed.

3.2.1. **Origins of the Analysis Task**

In the case of this study, the meaningful stimuli (fixed-frequency sinusoids) were embedded in a masking signal derived from either white noise or solar wind turbulence, such that auditory and visual performance might be assessed in the presence of varying levels of distractor stimuli. Additionally, the latter case closely resembles an analysis task that a heliospheric research scientist might encounter in the field, as these transient bursts of sinusoidal activity closely mirror several wave modes (e.g., whistler modes and ion cyclotron waves) that can be found in high-resolution magnetometer observations of solar wind turbulence. These waves are of interest to the scientific community because they effectively interact with particles; however, they are often very transient in nature and

difficult to identify through traditional analysis methods due to both the turbulent nature of the solar wind and the large volumes of available data.

An extremely clear example of one such event occurred in *Wind* magnetometer data during June 2008, which is displayed in **Figure 22**. Here, a spectrogram representation is presented that spans roughly 83,000 audified data samples derived from *Wind* magnetometer observations. Broadband turbulence is manifested as vertical lines, while wave activity is apparent as a single bright object at the center of the spectral display. This is one particularly clear example; most instances are subtle in comparison.

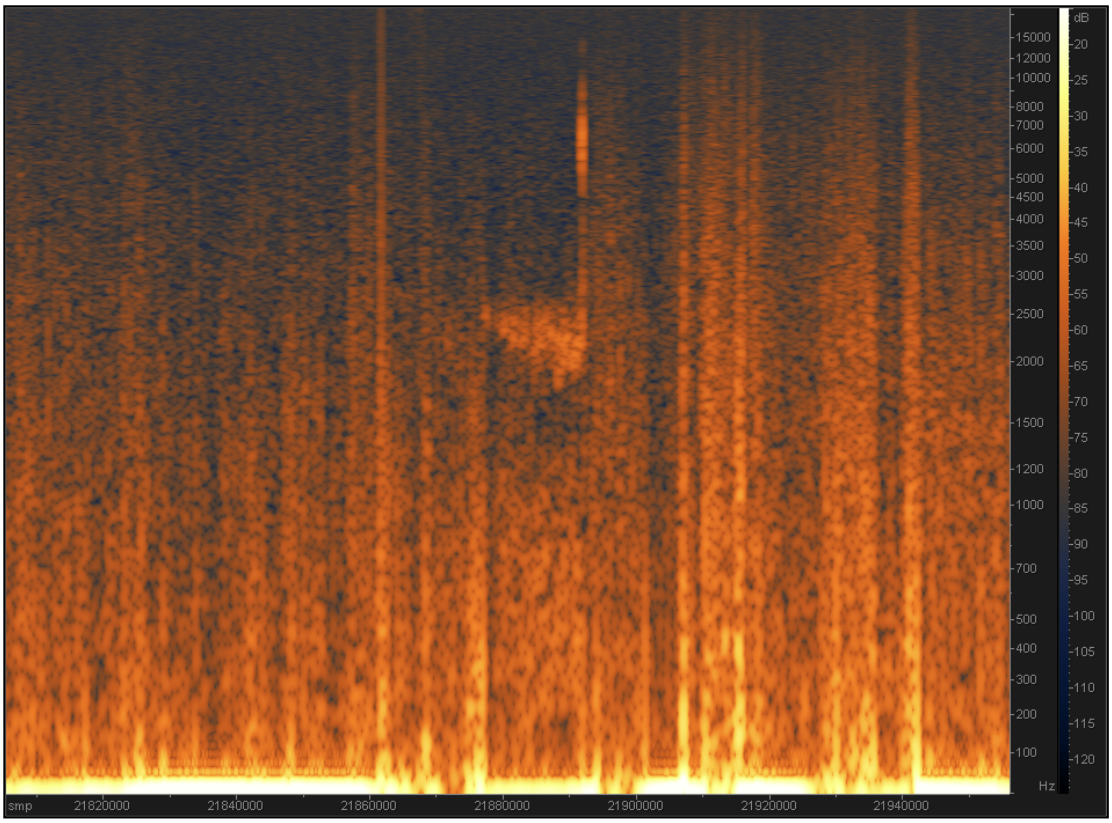


Figure 22. The spectrogram display (reduced in size) of a coherent wave event occurring in high-resolution *Wind* Magnetometer data during June 2008. This event spans roughly 23 minutes in the original data and 350 ms in the resulting audio file.

3.2.2. **Experimental Method**

One important guiding question has been: what baseline metrics can be established for auditory display through audification, and how do they compare to visual analysis capabilities? Toward this end, this study directly compares the analysis capabilities of participants who both listened to and viewed data as part of a structured

feature identification task. Transient sinusoidal waveforms were embedded in time-varying signals that contained broadband noise. The task of the participant was to identify how many of these transient events occurred within each example. The embedded sinusoidal events were tightly parameterized such that deeper investigation might provide some insight as to the performance of the two modalities in the identification of stimuli with varying amplitude, frequency, and duration. This kind of baseline evaluation is critical in order to gain a deeper understanding of how auditory perception may be applied to complex data analysis tasks, and ultimately integrated into the exploration of large data sets within the sciences.

3.2.2.1. **Hypothesis**

Participants will be able to identify the number of time-varying fixed-frequency sinusoid events more quickly and accurately through auditory data analysis than through visual analysis. Here, accuracy is a comparative measure of the number of events reported by the participant for each example versus the number of events that were actually embedded; this measure provides a margin of error.

3.2.2.2. **Participants**

Ten participants took part in this research study. Half were members of the Solar and Heliospheric Research Group (SHRG) at the University of Michigan, and the other half were computer-music specialists; all had experience working with spectrogram displays.

3.2.2.3. **Stimuli**

All solar wind data utilized in the study were gathered from magnetometer observations on the ACE and *Wind* satellites. These time series data sets were converted to audio files using an audification code written in Matlab. All data samples from the original data sets were preserved in this isomorphic mapping process. All visual stimuli were then rendered in the iZotope Rx software environment, and consistent settings were utilized to ensure uniformity across examples. Spectrograms were presented with a linear scaling on the y-axis, and a chromatic color mapping from black representing the absence of energy to white representing the full presence of spectral energy. This visualization

method was reviewed by members of the SHRG and deemed appropriate for the spectral representation of solar wind data. **Figure 23** is an example derived from measurements of the magnitude of the solar magnetic field as observed by the *Wind* spacecraft over a period spanning August to September 2004.

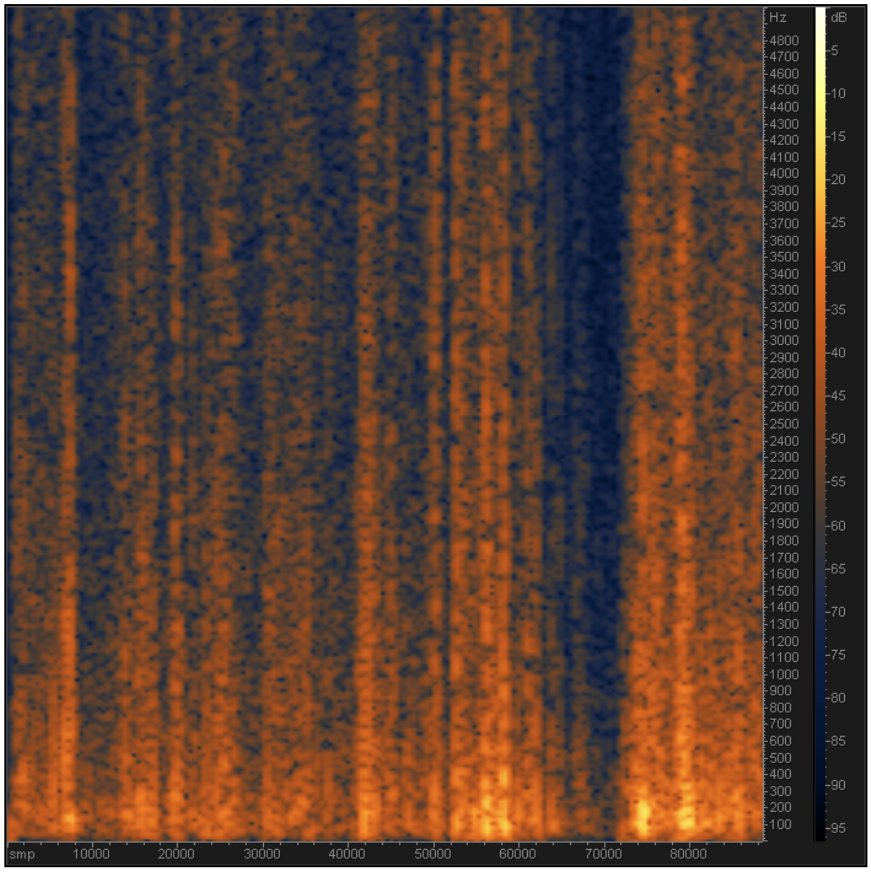


Figure 23. Spectral representation of audified solar wind turbulence. Broadband turbulence is represented as vertical bands of increased brightness.

The fixed-frequency sinusoid events were created with a synthetic data generation module constructed in the Max/MSP computer-music programming environment. These events ranged in frequency from 300 Hz to 4.7 kHz; intensity varied between -16 dB, -19 dB, and -22 dB (all masking noise was balanced to an RMS level of 0 db); and length varied between 25 ms, 50 ms, 100 ms, and 200 ms. The loudness level, frequency, and duration were held constant within each example, and varied between examples. The number of fixed-frequency events embedded in each example ranged from 0 to 3. All possible permutations of the data parameters were utilized to create a set of 48 unique stimuli that were embedded in both white noise and solar wind data, resulting in a total of

96 examples. These noise elements acted as masking signals with varying level of complexity. The solar wind data sets were pre-screened in order to minimize the likelihood that they inherently contained any significant fixed-frequency events that may be identified in the auditory and visual analysis tasks.

All examples contained approximately 88,000 data points, which translated to two seconds of audio playback at a sampling rate of 44.1 kHz. Four duplicate examples were included for the purpose of confirming internal consistency. All 100 examples were presented to both the auditory and visual modalities in an order that was randomized before the tasks began. The ordering of the two analysis tasks was also randomized across all participants, such that some completed the visual analysis module before moving onto the auditory analysis section, and vice-versa.

3.2.2.4. **Apparatus**

The experiment was conducted on a 15-inch MacBook Pro with the Mac OS X (10.8.2) operating system. The listening task was completed with Audio-Technica ATH-M50 stereo headphones. The pre-test, analysis tasks, and post-test were all encapsulated within a single standalone application constructed with the Max/MSP computer-music programming environment (version 6.05). All responses were recorded using the "coll" object and saved as data files in *.txt* format. A time-stamp for individual responses was recorded, along with total completion time for each task. Before beginning the experiment, participants were prompted to provide their first name, middle initial and last name; unique 3-letter file names were created from the initials.

3.2.2.5. **Procedure**

Participants were trained to visually and auditorily assess for the presence of fixed-frequency sinusoid events that were embedded in both white noise and noise generated from solar wind data sets. All visual stimuli were presented as spectrogram displays, and auditory stimuli were presented through audification and played back over headphones. These examples were presented sequentially, and participants were not allowed to go back and change their responses once an answer had been provided. The participants' task was to assess each example, and to report the number of fixed-frequency events they were able to detect. During the analysis task, participant responses

were entered into a number box that allowed any integer values between 0 and 99. These values could be entered either by clicking and dragging on the number box, or typing on the keyboard.

One training module guided participants through the process of listening to auditory data, and the other provided assistance in conducting a visual analysis of a spectrogram. These modules both explained the analysis task and guided participants through the interface (the visual training module is displayed in **Figure 24**). The data files used for the training sessions first demonstrated the fixed frequency sinusoids in isolation before introducing the full range of examples that participants would be expected to identify. These data examples were generated specifically for the training task, and were not included in the study. Additionally, participants were not able to change the volume setting once they completed the auditory training module; this prevented a perceptual bias that could be introduced by a global volume change partway through the experiment.

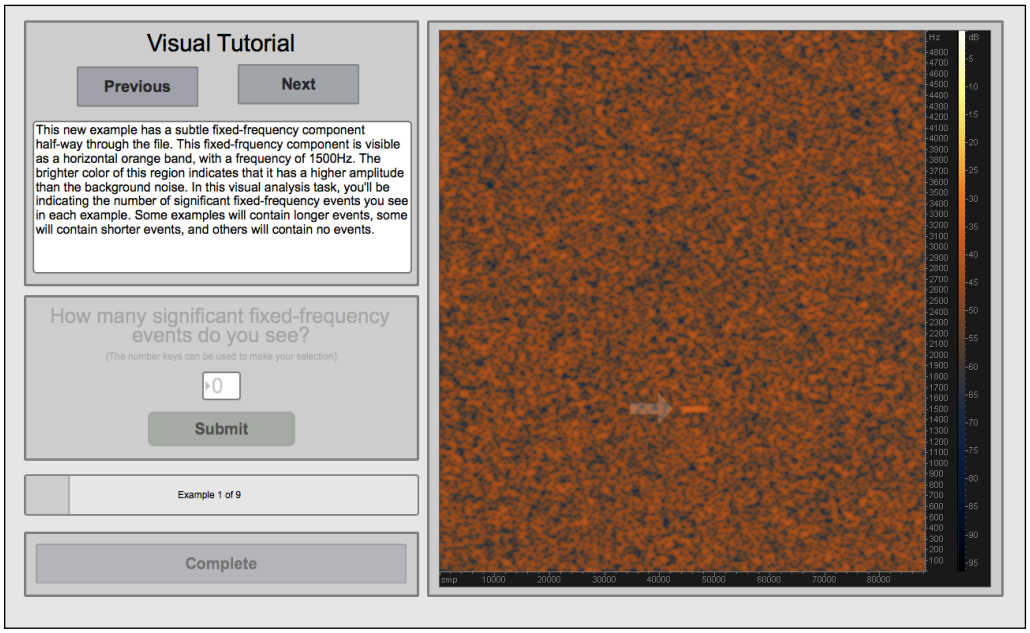


Figure 24. The visual training module provided participants with clear examples of what types of stimuli could be presented in the study.

An exact definition of "fixed-frequency" events was provided in training modules for both modalities, along with examples that demonstrated the types of events

participants would be expected to find. All participants reported that they considered the training provided for the analysis task to be easy to understand and/or adequate.

For the auditory portion of the task, participants were provided with several options for starting and stopping playback. An on-screen play-bar could be used to start the sample from any specific location, and the space bar could be used to start and stop playback. Additionally, a looping option allowed participants to listen to the audio repeatedly. For both the auditory and visual portion of the test, a small temporal gap was inserted between the presentation of each example in order to minimize the impact of subtle differences that may be present between subsequent stimuli.

Testing was conducted at various locations at the University of Michigan. All participants completed the analysis tasks in a quiet space that was free from potential distractions. The experiment was administered with an interface constructed in the Max/MSP computer-music programming environment. Subjects were provided with headphones and given a brief verbal overview of the task. After completing a short pre-test questionnaire, participants were randomly assigned the visual or auditory task. Participants were informed that while there was no time limit for this test, the total completion time was recorded, and they should attempt to respond "both quickly and accurately." In order to minimize the effects of fatigue, participants were informed that they could take a short break between the visual and auditory analysis tasks.

The post-test questionnaire was specifically designed to determine the participant's familiarity with sonification, experience working with spectrograms, level of comfort with computers, and experience with data analysis, mathematical modeling, and scientific research. Participants were asked to rate the difficulty of the listening task in relation to the visual task. This information was gathered in order to assess for a potential correlation between individual backgrounds and performance on the analysis task. Participants were also asked if they noticed any duplicate audio and/or visual stimuli, and if yes, to write how many. While the duplicate stimuli could be used as a measure of internal consistency, a participant's awareness (or lack thereof) of these duplicate stimuli could also yield potentially valuable insight as to their cognitive state during the examination. Finally, a space was provided for additional feedback in free-response form.

3.2.3. **Results**

In all instances, statistical significance was calculated through the implementation of a matched, 2 tailed *t*-test. For this study, significance was considered at a value of $p < .05$, and strong significance at a value of $p < .01$. Overall, participants provided correct responses for 66% of the visual stimuli, and 60% of the auditory stimuli, this difference of 6 percentage points was found to be statistically significant ($p < 0.01$). For examples in which fixed-frequency events were embedded in white noise, participants provided correct responses for 66% of the visual stimuli and 54% of the auditory stimuli ($p < 0.01$). For examples in which fixed-frequency events were embedded in noise generated from solar wind data sets, participants provided correct responses for 65% of the visual stimuli and 66% of the auditory stimuli ($p = .94$). A summary of task performance has been provided in **Table 5**, and additional information has been provided as to performance with the white noise and solar wind data maskers.

Table 5. A comparison of auditory and visual performance on the structured identification task. Here "SW" stands for solar wind.

		Vis	Aud	*p* value
Percentage of Correctly Identified Examples	White Noise	66	54	< 0.01
	SW Data	65	66	= 0.94
	Average	**66**	**60**	**< 0.01**
Average Error (Per Example)	White Noise	0.55	0.76	< 0.01
	SW Data	0.57	0.48	= 0.01
	Average	**0.56**	**0.62**	**= 0.03**
Total Number of Events Reported	White Noise	53	68	< 0.01
	SW Data	47	67	< 0.01
	Total	**100**	**135**	**< 0.01**
Average Completion Time (mm:ss)	**Average**	**11:45**	**18:00**	**< 0.01**

The margin of error was calculated by first determining the difference between the participant responses and the correct response for each example, and then averaging across the total number of examples. Participants had an average error margin of 0.56 in the visual analysis task, and 0.62 in the auditory analysis task ($p = 0.03$). For examples in

which fixed-frequency events were embedded in white noise, participants had an error margin of 0.55 in the visual task and 0.76 in the auditory task ($p < 0.01$). For examples in which fixed-frequency events were embedded in noise generated from solar wind data sets, participants had an error margin of 0.57 for the visual stimuli and 0.48 for the auditory stimuli ($p = 0.01$).

The average completion time on the visual analysis task was 11 minutes and 45 seconds; this was 53% faster than the average completion time for the auditory task, which was approximately 18 minutes. On average, participants reported detection of 100 total events in the visual analysis task and 135 events in the auditory analysis task (a 35% increase).

A summary for overall task performance as a function of the number of embedded events is provided in **Figure 25.** Participants utilizing visual analysis correctly identified examples without any embedded stimuli at an average success rate of 96%, this rate was 71% in the auditory analysis task. Respectively, average rates for the successful identification of single events were 53% and 50%; double event identification rates were 53% and 58%; and successful triple event identification rates were 62% and 60%. The difference in performance between the two modalities was only statistically significant for examples containing no events ($p < 0.01$).

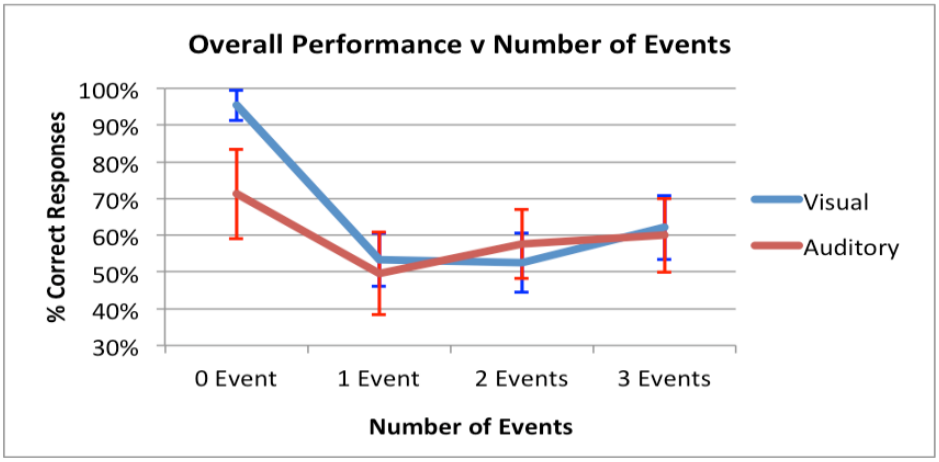


Figure 25. Task performance as a function of the number of embedded fixed-frequency events. Error bars indicate standard deviation.

A summary for the overall task performance as a function of the intensity of embedded events is provided in **Figure 26**. Performance generally declined as the intensity of embedded events decreased. Events provided at -16 dB were visually

identified with a success rate of 83%, and auditorily identified at a success rate of 80%. These respective values for events that were presented at -19 dB were 46% and 48%; for events presented at -22 db, successful identification rates dropped to 38% and 39%. These differences were not found to be statistically significant.

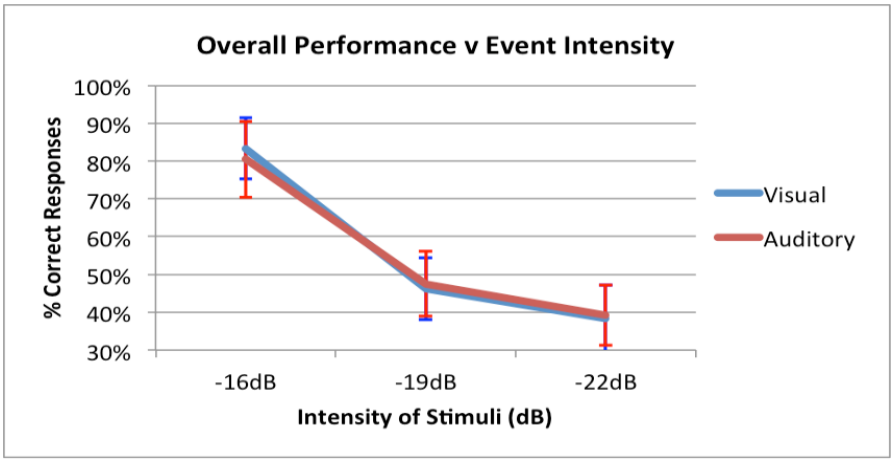


Figure 26. Task performance as a function of stimuli intensity. Error bars indicate standard deviation.

A summary of the overall task performance as a function of the length of embedded events is provided in **Figure 27**. Stimuli that contained embedded events with a duration of 200ms were correctly identified visually at an average success rate of 76%, and correctly identified auditorily at an average success rate of 74%. The respective success rates for stimuli containing events with a duration of 100ms were 75% and 73%; success rates dropped to 72% and 64% for events lasting 50 ms in duration ($p = 0.018$), and 45% and 53% for events with a duration of 25 ms ($p = .049$).

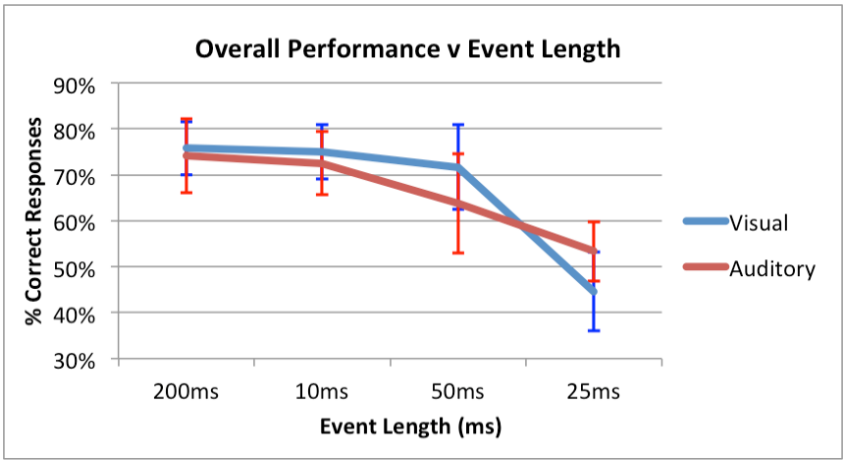


Figure 27. Performance on the identification task as a function of decreasing event length. Error bars indicate standard deviation.

3.2.4. **Discussion**

When performance across all 100 examples is assessed, vision outperformed audition by a margin that was statistically significant ($p < .01$). However, the auditory modality consistently outperformed the visual modality in the detection of fixed-frequency sinusoid events embedded in solar wind data sets ($p = .014$). In this case, the average visual identification success rate was 52%, while the auditory success rate was 63%. **Figure 28** provides a summary of the performances of the two modalities in identifying events embedded in solar wind data sets at various levels of intensity. Here it can be seen that the success rate for auditory recognition improved slightly in relation to visual recognition as event intensity declined. Participants who utilized the auditory modality correctly identified 50% of examples that contained events embedded at -22 dB, while participants who utilized visual analysis successfully identified 40%. This difference of 10 percentage points was found to be statistically significant ($p = 0.01$).

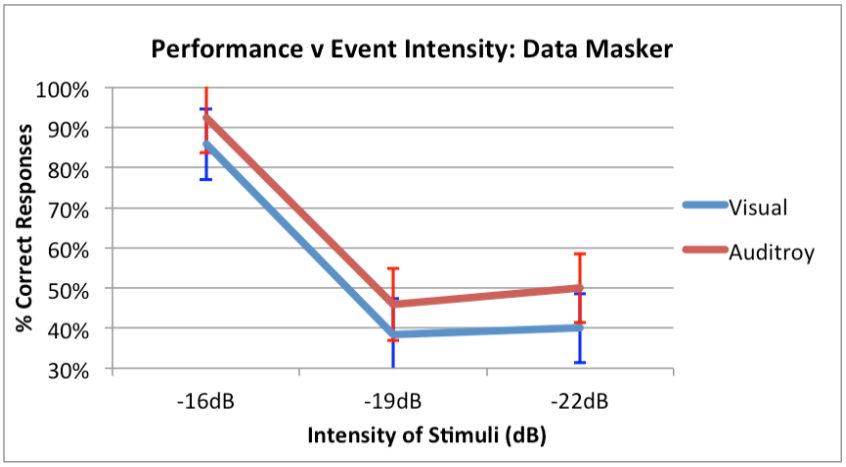


Figure 28. Percentage of correctly identified examples containing fixed-frequency events embedded in solar wind data as a function of the intensity of embedded events. Error bars indicate standard deviation.

3.2.4.1. **Analysis of Demographic Influence**

Demographic information and previous experience (as assessed by the pre- and post-tests) significantly contributed to task performance in many instances. Correlation between demographic information and task performance was determined by calculating Pearson's product-moment correlation coefficient [198]. Equation (1) is used to calculate linear dependencies between two variables, where r is the sample correlation coefficient and x and y are the variables under test. A resulting correlation factor of $r = 1$ indicates a perfect correlation between the two variables, a factor close to zero indicates very little or no correlation, while a factor of -1 indicates a perfect anti-correlation.

A strong positive correlation was found between task performance and educational level, with performance increasing as a function of number of years in higher education ($r = 0.83$). A moderate correlation was found between successful identification rates on the two tasks ($r = 0.63$). A moderate negative correlation was found between performance on the two tasks and amount of musical training ($r = -0.46$), as well as a moderate negative correlation between task performance and experience with sound editing and audio processing ($r = -0.66$). One participant found the visual analysis task to be more difficult than the auditory task; five considered the identification task to be easier. Perceived difficulty had no statistically significant effect on task performance.

3.2.4.2. **False Positives in the Auditory Identification Task**

While the visual modality had a higher overall success rate in identifying the number of embedded fixed-frequency components, the data reveal several pieces of insightful information upon closer inspection. When the stimuli containing zero events are removed from the analysis, the overall success rate in both modalities evens out at 56%. It is immediately clear that the auditory modality was predisposed toward indicating false-positives in the absence of embedded frequency components. The distribution of these false positives was not completely uniform, multiple examples were labeled as containing one or more embedded events by half of the participants, while other examples were either correctly identified by all participants or incorrectly identified as a false positive by a single participant. White noise is not a truly randomized distribution, and it is possible that upon repeated listening, some participants began to pick up on structures occurring at very small time scales.

One participant noted that they attempted to carefully fine-tune their auditory threshold for event detection, and that they "didn't include some short frequency bursts that may be audible." This suggests that they were indeed able to hear short transient fixed-frequency events occurring more rapidly than the 25 ms threshold. Another participant noted that they "heard more happening (in the audio file)… This was good but also leads me to wonder if I had some false positives." The eye had comparatively little trouble detecting events presented in the relatively uniform visual background created by white noise—the non-uniform broadband energy bursts present in the solar wind data sets may have provided a significant amount of visual distraction.

3.2.4.3. **Closer Investigation of Individual Stimuli**

There were several examples utilizing solar wind data that were consistently assessed incorrectly through visual analysis, and correctly assessed through audition. One such example was missed by every participant in the visual task and only missed by one participant auditorily. This example contained a 25 ms event at -22 dB that occurred very close to the end of the data set. Another example that contained a slightly longer event very close to the end of the file was correctly identified only once visually, and eight

times auditorily. This indicates that there may potentially be some visual bias away from events that take place at the very edge of a spectrogram.

One example was incorrectly identified seven times visually and only once auditorily. The fixed-frequency element occurring roughly half way through the file at 1.5 kHz was subtle, but easily recognized auditorily with training. A second fixed-frequency component, beginning half way through the file, was almost completely visually obscured by the broadband noise event that occurred at the same time. While the same event was also masked by the broadband noise element in the auditory representation, it seems that the ear may not have had as much trouble separating the sinusoidal signal from the background noise.

3.2.4.4. **Review of Experimental Design**

This study utilized a relatively small pool of ten participants. The recruitment of a larger participant pool was hindered due to the lack of available individuals with the specialized knowledge necessary to complete the identification task. Ideally, future research should work with a larger sample size in order to better determine statistical significance. The use of participants with domain-specific knowledge limits the transferability of these results, as the performance may vary in the general public.

3.2.4.5. **Evaluating Internal Consistency**

For each participant, four examples were presented twice in each modality in order to determine whether participants' responses were consistent across multiple exposures (the majority of participants indicated that they detected the presence of repeated stimuli when asked in the post-test). No participant was entirely consistent across the four repeated stimuli, and on average participants were consistent in their evaluation of approximately 3 of the 4 stimuli for both modalities. Three participants answered consistently across all repeated visual examples, and three separate participants achieved perfect consistency auditorily. This lack of complete internal consistency indicates that participant evaluations varied slightly over time, which could be attributed to factors such as learning or fatigue, which might improve or degrade performance over time respectively.

This speaks to the difficulty of the analysis task, which required participants to assess for the presence of extremely subtle features. In light of this fact, the lack of complete internal consistency is to be expected, and the effects of learning and fatigue were minimized through both the randomization of the task ordering across participants, and the randomization of the stimuli presented within these tasks.

3.2.4.6. **Discussion**

Though both identification tasks engaged separate modalities, they were fundamentally similar in that each involved the identification of pre-defined objects embedded within background noise. Visually, these objects were defined by dimensions of color, brightness, length, width, and height; while auditorily they were defined by the frequency space they inhabited, their relative amplitude, and duration. Placed within the context of gestalt theory [199], it could be said that participants utilized these unique properties in establishing, for example, "belongingness" for an explicit subset of the incoming sensory stream. The results of this study, generally speaking, provide some information about the relative ability of the visual and auditory modalities to segregate meaningful information from background noise in the evaluation of certain scientific data sets.

Visually, the spectrogram display of white noise presented a relatively uniform background characterized by a lack of remarkable structures, and this visually unified pattern could be perceptually encoded as a single object against which the embedded features were readily identifiable. Conversely, the spectrogram display of the solar wind spectra contained features on both micro- and macro-scales. The difference between visual and auditory performance when features were embedded in solar wind spectra suggests that the visual modality was comparatively more affected by the presence of complex distractor stimuli than audition, and that this effect was greater in the identification of subtle features. Additionally, the discrepancy between visual and auditory performance in the identification of stimuli embedded in a synthetic noise mask suggests that future research should test the original hypothesis with noise other than white (e.g., pink).

It could be said that the mechanisms that promote auditory stream segregation were brought to bear as participants listened to sounds derived from solar wind data sets, and the auditory system was comparatively more successful in using subtle spectral cues to parse meaningful signals from background noise [65]. This points toward the types of features that may be best suited for recognition through audification—namely those which subtly present themselves within a complex time-varying signal. While visual performance surpassed audition in the identification of sinusoidal events that were 50 ms in length, auditory analysis yielded a higher success rate in the identification of the shortest events.

It is common practice for many heliospheric scientists to create visual representations that average a power spectrum over a large number of data samples, and in these instances audification could provide new information regarding the small-scale features that are lost in this process. In this type of practical data analysis task the strengths of one modality may support the weaknesses of another, as audification may reveal subtle spectral features overlooked through visual assessment, while vision may assist in ruling out events that are too subtle to warrant additional investigation.

A future study should investigate the ability of the auditory and visual modalities when applied in tandem toward a specific scientific data analysis task. In this way, some insight could be gleaned as to how audification compares with visual analysis techniques in real-world scenarios. Additionally, this would shed light on the types of features that are readily identifiable through auditory analysis. An interface such as iZotope Rx is an ideal platform for conducting such work, as it provides real time feedback both visually and auditorily, and annotations may be added directly to the data in the form of markers. While this study employed highly parameterized artificially generated stimuli in order to extract some quantitative information as to the relative performance of the two modalities, future research should draw example stimuli from raw data sets as found in the field, and participants could be provided with a more open-ended identification task.

Finally, it is worth noting that, for many of the participants, this was the first instance in which they had utilized auditory analysis in a data analysis task, while some had worked with spectrogram displays for well over a decade. For this reason, it would

also be valuable to study the effects of training on participant performance in a set of structured analysis tasks.

3.2.4.7. **Additional Analysis**

This section contains results that were not included in the original publication. It was noted that a *p* value was not calculated for the disparity in completion time between the two modalities—the difference of 53% was found to have strong significance ($p < .001$). A larger disparity was found in the study presented in Chapter VII, and here it was suggested that participants' lack of exposure to auditory methods likely impacted completion times in the auditory condition. For this reason the disparity was attributed to the novelty effect, and increased exposure to auditory methods was recommended through a training environment.

The results from the auditory analysis task with fixed-frequency components embedded in white noise are plotted as a function of event frequency In **Figure 29**. At first glance it appears that contour of the curve follows the equal-loudness measurements of Fletcher and Munson presented in the second chapter. In these data it appears that the area of highest sensitivity occurred at 1.9 kHz, and performance sharply decreased above 2 kHz (offset somewhat from Fletcher and Munson's original curve). It would be expected that identification rates would be highest in the range of maximum auditory sensitivity—for this reason researchers should explore new data sets at a variety of playback speeds in order to assess for the presence of subtle events in a variety of frequency regimes. Future research should more closely examine this effect in an applied analysis task.

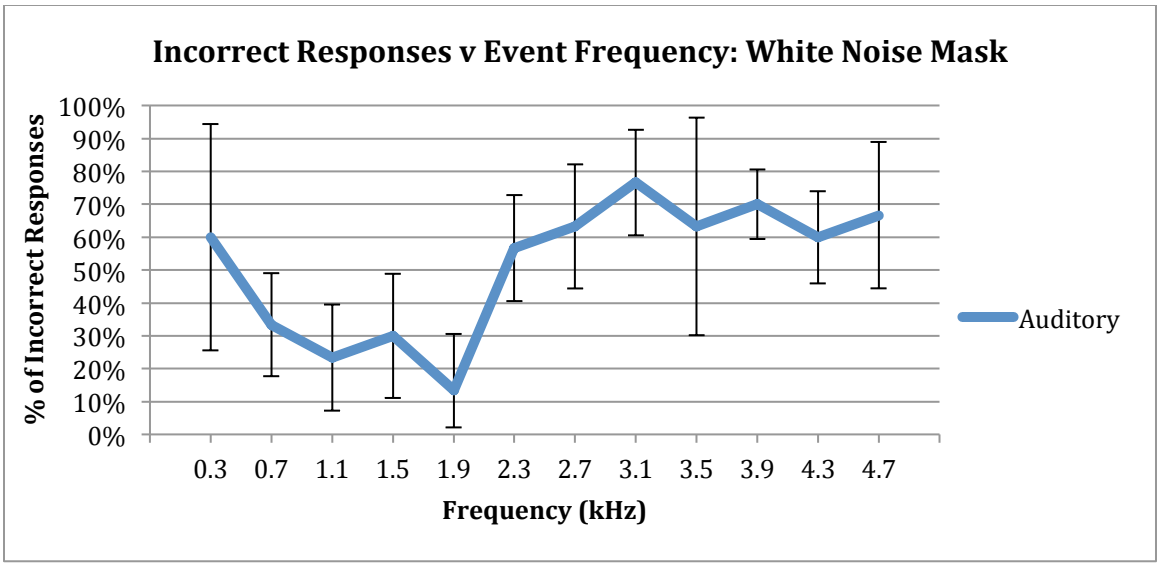


Figure 29. Identification of fixed-frequency components in white noise as a function of frequency. Error bars indicate standard deviation.

3.2.5. **Conclusion**

This study directly compared the performance of participants utilizing auditory and visual analysis methods in a structured data analysis task. While visual analysis outperformed auditory analysis overall, additional investigation revealed that auditory analysis outperformed vision when events were embedded in solar wind data as opposed to white noise. In these instances, the identification task closely resembled a type of analysis conducted by heliospheric scientists. When provided with examples that contained no embedded fixed-frequency events, participants utilizing the auditory modality were more likely to report false positives, and it was suggested that this could be attributed to the extreme sensitivity of the auditory modality. Finally, several data examples contained embedded events that were correctly identified through audition while consistently overlooked through visual inspection. These findings support earlier research that revealed a high correlation between assessments made through auditory and visual analysis methods, and further suggest that the analysis capabilities of each modality may vary based largely on the complexity of the masking stimuli that are present.

3.3. Auditory and Visual Identification of Anomalous Features in Solar Wind Time Series Data

The previous two studies begin to establish a set of baseline metrics for comparing auditory and visual data analysis capabilities. In order to bring the larger picture into focus a slightly more open-ended question may be posed—can the nuances of the classification schema employed by both modalities be uncovered in a task that emulates a real-world analysis scenario? Gestalt psychology provides a means to operationalize this inquiry by framing data analysis as sense-making tasks and feature identification as a process of perceptual grouping [65, 200]. The previous study adopted a bottom-up perspective, using tightly controlled stimuli to evaluate and compare perceptual thresholds across the two modalities. This study incorporates a top-down approach—requiring research scientists to categorize features observed in solar wind time series into a set of clearly defined groups, and to label any features deemed to be anomalous.

This task sheds light on how the perceptual grouping mechanisms of the two modalities influence the establishment of "belongingness" in the exploration of solar wind time series. In this sense, features are divorced from their scientific context and treated purely as auditory and visual objects. This study, then, seeks to qualitatively and quantitatively evaluate how these objects are classified in the exploratory analysis of heliospheric data sets, and to investigate whether auditory analysis (through audification) may result in the identification of a higher number of anomalous features. To this end, participants were provided with two data examples gathered by the MFI instrument on the *Wind* spacecraft [201], and two examples from the magnetometer instrument on STEREO [202]. Drawing from active missions, this study provides an exploratory data analysis task that closely mirrors those that may be encountered in the field.

3.3.1. Experimental Method

3.3.1.1. Hypothesis

Using audification, participants will be able to more rapidly and successfully identify anomalous events that are characterized by temporal patterns unfolding in data

streams derived from one-dimensional time series.

3.3.1.2. **Participants**

Eight participants were recruited for this study, all of whom worked as full-time researchers in the Heliophysics division at NASA Goddard Space Flight Center (GSFC). All participants had some familiarity in working with spectrogram displays. The group included 5 males and 3 females, and ages ranged from 24 to 70 with an average age of 45. In a self-assessment, the majority of participants reported average to above average hearing, while no participants indicated that their hearing was poor. Half the participants had 7 or more years of musical training, two had received 2 to 6 years of musical training, and two had none. All participants held Ph.D.'s except for one, who was working towards a master's degree.

3.3.1.3. **Stimuli**

A total of four solar wind data sets were audified as 16-bit files in .aif format using an audification code written in Matlab. These files were comprised of two time regions selected from data gathered by the MFI instrument on the *Wind* spacecraft [201, 202]. These data sets ranged in length from approximately 600-thousand samples to 2-million samples, with an average length of 1.3-million samples. The corresponding audio files ranged in length from 20 seconds to 47 seconds, with an average length of 35 seconds.

Both data examples from the *Wind* spacecraft were taken from the *z* component in the GSE coordinate system, which points north out of the plane of the earth's orbit around the sun. This component was chosen due to the relatively low presence of instrumentally induced noise produced by the rotation of the spacecraft (in comparison to the *x* and *y* components). The first *Wind* example spanned 1,335,442 data samples from magnetic field measurements gathered on the 22nd and 23rd of January, 2010. The resulting audio file was approximately 30 seconds in length at a playback rate of 44,100 samples per second. The second example from the *Wind* spacecraft spanned 1,333,872 data samples from magnetic field measurements gathered on the 12th and 13th of November, 2010. The resulting audio file was approximately 30 seconds in length at a playback rate of 44,100 samples per second.

The first STEREO example spanned 1,324,995 data samples from magnetic field measurements gathered between January 23rd and 26th, 2007 (by the STEREO A spacecraft). The resulting audio file was approximately 30 seconds in length at a playback rate of 44,100 samples per second. The measurements were taken from the *y* component in the GSE coordinate system, which points backward behind the earth as it orbits the sun.

The second STEREO example spanned 2,073,517 data samples from magnetic field measurements gathered on the 10th, 11th, and 12th of November 2010 (by the STEREO B satellite). The resulting audio file was approximately 47 seconds in length at a playback rate of 44,100 samples per second. The measurements were taken from the *z* component in the GSE coordinate system.

3.3.1.4. **Apparatus**

The study was administered on a 15-in Macbook Pro with the Mac OS X 10.8.5 operating system. All listening tasks were completed with Audio Technica ATH-M50 professional studio monitor headphones. The analysis task was conducted using the iZotope Rx 2 software platform (version 2.10.656). All sessions took place at NASA GSFC in a quiet space that was free of distractions. While the majority of participants completed the entire study in one session, a few participants required multiple sessions to complete the analysis task for both modalities. All instructions and training modules were printed and provided to the participants at the beginning of the study. The post-test was administered as a standalone application constructed within the Max/MSP computer-music programming environment (Version 6.1.3). All responses were recorded using the "coll" object and saved as data files in *.txt* format. At the beginning of the post-test, participants were prompted to provide their first name, middle initial and last name; unique 3-letter file names were created from the initials.

3.3.1.5. **Procedure**

The pilot-study presented here incorporates a within-subjects experimental design in which participants were asked to conduct an assessment of audified solar wind data sets using the iZotope Rx software package. Participants were presented with two analysis tasks in a randomized order: visual analysis of a spectrogram display, and

auditory analysis of audified data (a total of four data examples were provided). Two data sets were presented visually as a spectrogram display with an overlaid line-plot of the time series data; two data sets were provided auditorily, and audio playback was provided in lieu of a spectrogram display. Participants were trained in the use of iZotope Rx to assess for spectral features in a digitally manufactured data example; this process involved **1)** the identification of groups of features (both auditory and visual) by marking the best representative example and assigning a name, **2)** the identification of features within a group that appeared to be anomalous, and **3)** the identification of anomalous features that did not appear to belong to any group.

In order to evaluate participant performance, these results are compared against those of an expert who completed the same grouping and anomalous feature identification task. The expert is a research scientist at NASA GSFC with several years of experience working with high-resolution magnetometer data and deep domain-specific knowledge of structural features found in the time series produced by the magnetometer instruments on the *Wind* and STEREO spacecraft.

3.3.2. **Results**

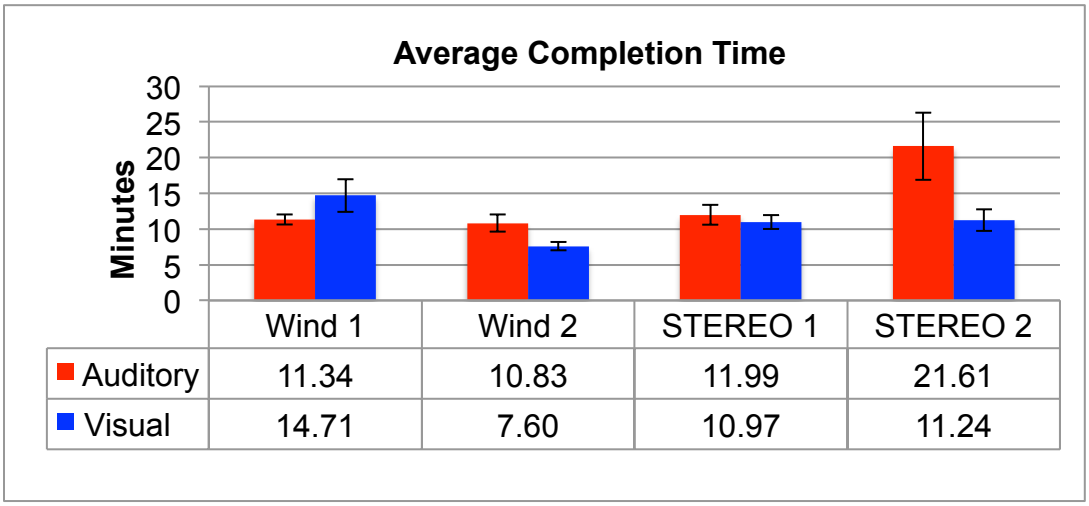
In all instances, statistical significance was calculated through the implementation of a two-tailed *t*-test. A matched *t*-test was used when comparing a participant's performance across the auditory and visual modalities, and an unmatched *t*-test was implemented when calculating performance within individual data examples (as no participant was presented with a single example in both modalities). For this study, significance was considered at a value of $p < .05$, and strong significance at $p < .01$.

3.3.2.1. **Completion Time**

The average completion time across all examples in the auditory task was 13 ± 5.1 min, while the average completion time in the visual task was 11 ± 2.9 min, a difference of approximately 25% in additional time ($p = 0.35$).[25] The average completion time for each data example is displayed in **Figure 30**. The visual assessment was completed more rapidly than the auditory assessment for all examples except for the first

[25] Three participants completed the auditory task more quickly (on average), while the remaining 5 were faster in the visual task.

Wind data sample. The respective average completion times for auditory and visual assessment were approximately 11 ± .69 min and 14 ± 2.3 min for the first *Wind* Example, 11 ± 1.2 min and 8 ± 0.6 min for the second *Wind* Example, 12 ± 1.4 min and 11 ± 1.0 min for the first STEREO Example, and 22 ± 4.7 min and 11 ± 1.5 min for the second STEREO example. The difference in completion times across the two modalities was not found to be statistically significant for any data Examples.[26]



	Wind 1	Wind 2	STEREO 1	STEREO 2
Auditory	11.34	10.83	11.99	21.61
Visual	14.71	7.60	10.97	11.24

Figure 30. The difference in completion times between the two modalities was smallest in the first STEREO Example, and largest in the second STEREO Example. Error bars indicate standard deviation.

3.3.3. **Anomalous Feature Identification**

A scoring system was implemented to compare participant performance against expert assessment across the four data Examples. A value of "1" was assigned to instances in which a participant conclusively labeled the same feature as the expert, a value of "0.5" indicated an inconclusive but likely agreement between the two markings, and a value of "0" was provided for instances in which a participant failed to label a feature that was identified by the expert. The participant scores were then divided by the

[26] In the case of the second STEREO data Example, one participant's completion time was approximately three times slower than the average completion time for the auditory analysis task. It is for this reason that the standard deviation is quite large. If the data from this participant were removed, the timing results would be nearly identical across the two conditions. However, this value was not considered an outlier as the results of the analysis task for this participant were well formed, they successfully completed the grouping tasks provided in the training module, and reported full understanding of the analysis task (in the post test). As this was the first data example provided to the participant for auditory analysis, this timing disparity may be attributed to the novelty effect.

total number of features observed by the expert for each data example to determine the overall percentage of correctly identified features. This metric was applied to all instances in which the expert identified an anomalous feature (both within-group and without-group) in order to quantitatively evaluate participants' ability to identify anomalous features across the two conditions.[27]

The expert identified a total of eight anomalous features across the four data examples—seven purely anomalous features and one within-group anomaly. Seven of these features (86%) were missed by all participants in the visual analysis task, while on average every feature was identified by more than 50% of participants through the application of auditory analysis. Participants in the auditory condition displayed an average identification rate of 52% ± 16%, while participants using visual analysis displayed an average identification rate of 13% ± 0% (this small error margin is due to the nature of participant assessments in the visual condition—all participants visually missed seven anomalous features and properly identified one); these results are displayed in **Figure 31**. This difference of approximately 400% was found to be statistically significant ($p = 0.003$).

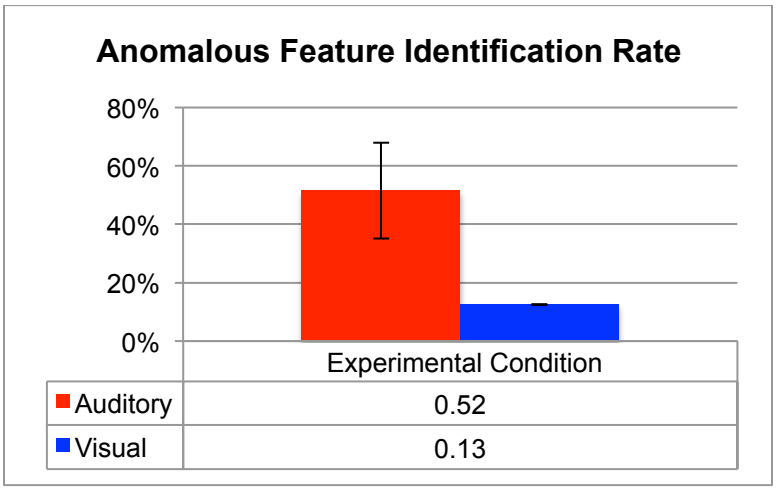


Figure 31. Participants using auditory analysis displayed a significantly higher level of agreement in the identification of anomalous features as compared to visual analysis. Error bars indicate standard deviation.

[27] Two events were identified by the expert as anomalous due to the presence of a "gap" in solar wind activity, these were excluded from the analysis presented here as they do not meet the criteria of an unfolding temporal event as stated in the hypothesis, and are not considered a "feature" as such.

3.3.4. **Discussion**

Through the application of auditory analysis, participants were able to identify more anomalous features in time series generated by magnetic field measurements gathered by the *Wind* and STEREO spacecraft, and this increased rate of feature recognition did not come with any notable impact on completion time. This discussion will investigate correlative effects in relation to demographic information that was gathered during a post-test, provide an assessment for participant agreement across all four data examples, present several anomalous features and discuss rationale for why they were or were not identified through visual and auditory assessment, explore the terminology applied by participants in labeling spectral features, provide additional details as to the nature of groups created by participants, assess the potential impact of the audio waveform on the results of the auditory analysis task, review subjective feedback provided by participants, and finally—present limitations and future avenues of research.

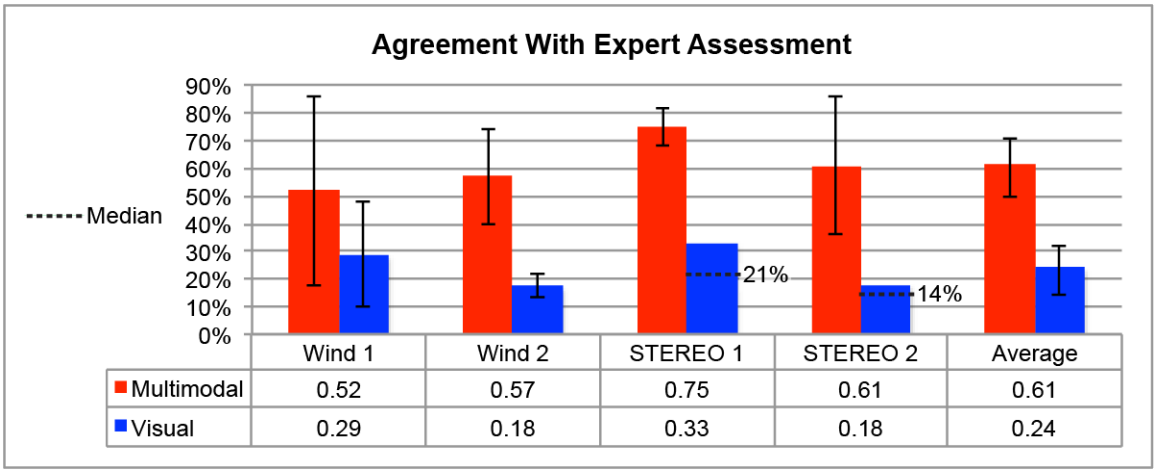
3.3.5. **Analysis of Demographic Influence**

This section will explore the potential correlative relationship between participant performance and information gathered during the post-test; in all instances correlation is determined through the calculation of Pearson's product-moment correlation coefficient [198]. No correlation was found between performance on either task and education level or gender. A moderate positive correlation was found between participant performance across the two conditions ($r = 0.41$), indicating some general level of analysis ability that may be translated across modalities. A small positive correlation was found between age and performance on the visual task ($r = 0.35$) while a small negative correlation was found between age and performance on the auditory task ($r = -0.29$). A moderate positive correlation was found between years of musical training and performance on the auditory analysis task ($r = 0.46$), though musical training had no correlation with performance on the visual analysis task. A very small negative correlation was found between participants' familiarity with spectrogram displays and performance on the auditory task ($r = -0.1$), while a moderate positive correlation was found between familiarity and performance on the visual task ($r = 0.46$). No correlation was found between perceived difficulty and performance on the task.

The strongest positive correlation (r = 0.77) was found between performance on the auditory task and the belief that sonification can reveal new insight in the investigation of scientific data sets. Participants who indicated this belief had, on average, a 63% higher agreement rate with expert assessment in the auditory analysis task. All participants who reported skepticism as to the potential usefulness of auditory analysis also described the listening task as more difficult than the visual analysis task. While no general correlation was found between perceived difficulty and task performance, the lowest agreement scores for the auditory condition were returned by participants who both found the listening task to be difficult and did not believe sonification could be useful.

3.3.6. **Expert Comparison**

The same metric that was used to compare anomalous regions identified by the expert with those identified by the participants can be extended to incorporate the classification of groups. In this way, expert and participant performances on the grouping task can be compared on the whole. **Figure 32** presents the results of this analysis for the four individual data examples along with the overall agreement rate for the two conditions.



	Wind 1	Wind 2	STEREO 1	STEREO 2	Average
Multimodal	0.52	0.57	0.75	0.61	0.61
Visual	0.29	0.18	0.33	0.18	0.24

Figure 32. In all data examples, participant-identification more closely aligned with expert-identification through auditory analysis. Error bars indicate standard deviation. The median value is provided in lieu of error bars for instances in which the distribution is heavily skewed by an outlier (in these instances the small number of valid data points leads to a highly non-Gaussian distribution).

For the first *Wind* data example, the average agreement was 52% ± 24% through auditory assessment and 29% ± 16% through visual assessment, a difference of 23 percentage points ($p = 0.11$). For the second *Wind* data example the average auditory agreement was 57% ± 21% and the average visual agreement was 18% ± 6%, a difference of 39 percentage points ($p = 0.007$). For the first STEREO data example the average agreement through auditory analysis was 75% ± 7% and the average agreement through visual analysis was 33% ± 28%, a difference of 42 percentage points ($p = 0.028$). For the second STEREO data example the average agreement through auditory analysis was 61% ± 21% and the average agreement through visual analysis was 18% ± 18%, a difference of 43 percentage points ($p = 0.02$). Overall, the average agreement between participant and expert assessments across all data examples was 61% ± 9% through auditory analysis and 24% ± 7% through visual analysis, this difference of 37 percentage points was found to be strongly statistically significant ($p < 0.001$).

3.3.7. **Investigating the Nature of Anomalous Features**

Only one anomalous feature (as identified by the expert) was uncovered by participants in the visual analysis task, and this feature was also identified by several participants through auditory analysis. Conversely, seven features were identified through auditory analysis that were overlooked by all participants in the visual analysis task. This section will explore the spectral characteristics and temporal evolution of several of these features in detail.

The expert identified four groups of features and two anomalies within the first data example drawn from the STEREO archive. The analysis presented in the previous section found this example to have the highest expert-participant agreement rate across both the auditory and visual conditions. This may be attributed to the fact that the data example is highly featured in comparison to the other three—containing distinct bursts of high-frequency activity, regions of increased spectral power at lower frequencies, as well as one anomalous feature identified by the expert toward the end of the file. This feature, situated between two periods of relative inactivity, contained a band of increased spectral power between approximately 300 and 800 Hz in the audification. Closer examination of

the temporal evolution of this feature reveals a steady rise in frequency before the activity comes to a relatively abrupt end (**see Figure 33**).

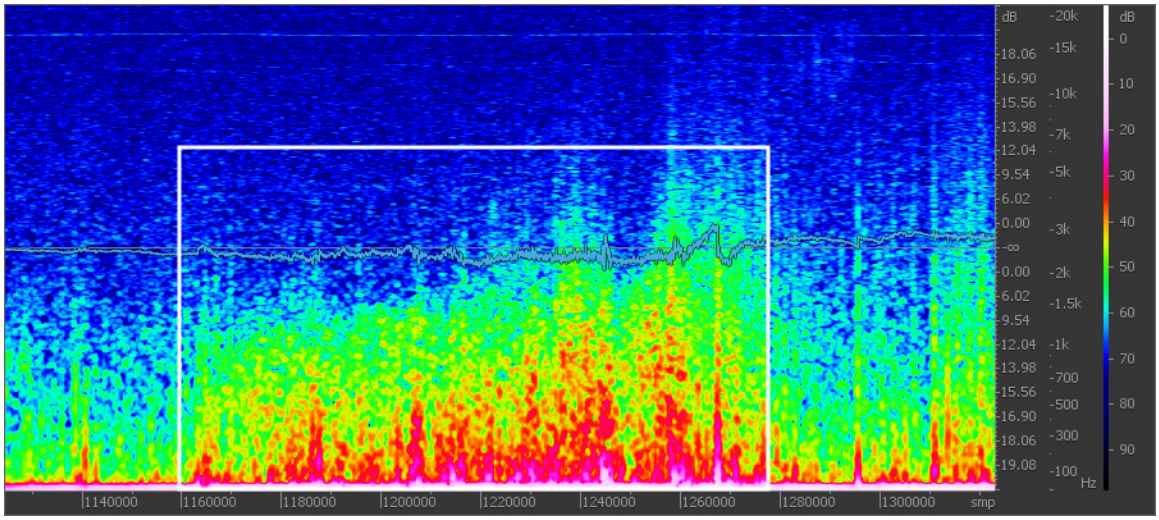


Figure 33. An anomalous feature (see white box) identified by the expert in the first data example from the STEREO satellite (shown here in iZotope Rx). The feature is located in the central 3/5^{ths} of the spectrogram display.

This feature was not classified as anomalous by any participants in the visual analysis task, but was identified in every auditory trial—one participant labeled this region as containing "chorus," while another noted that it sounded like a "storm outside." These participants were picking up on the subtle enhancement in mid-frequency activity that is not immediately apparent through visual inspection. The expert hypothesized that this feature was potentially caused by a *slow wind transient*, which he described as "a blob of plasma that has come from close to the solar equator, and was trapped close to the Sun before being released by a reconnection event." It was suggested that this event may have been generated by various populations of plasma mixing at the edge of the heliospheric current sheet. The associated data from the particle detectors on the STEREO A satellite were unavailable, hence a definitive classification for this event could not be provided.

The first data example from the *Wind* satellite contained another anomalous feature that was overlooked by all participants in the visual analysis task and regularly identified through auditory scan. Observing the area encompassed by the white box in **Figure 34,** it is immediately apparent why this feature may have been considered unremarkable through visual scan—while the central area displays a subtle increase in

spectral power (a small patch of yellow/red) it is relatively unpronounced and quite similar to other features found at lower frequencies.

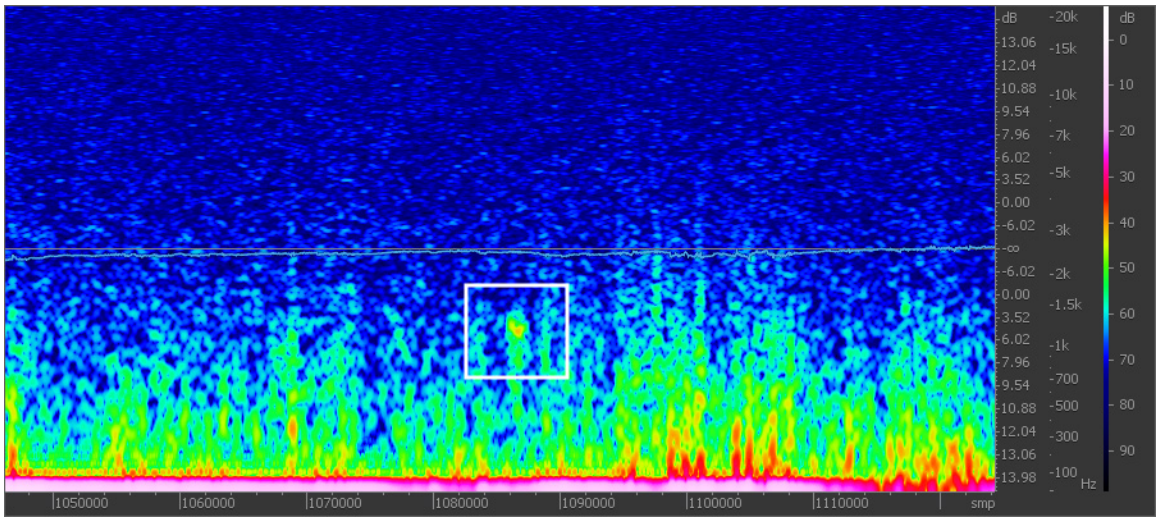


Figure 34. An anomalous feature (see white box) identified by the expert in the first data example from the *Wind* satellite (shown here in iZotope Rx). This feature was regularly overlooked by participants through visual analysis.

Closer investigation of the time series during this period reveals an interval of periodic oscillations spanning approximately 900 data samples in length, and lasting only 20 ms in the audification. Through auditory assessment one participant likened this feature to a "computer tick" while another noted that the sound resembled a "faint beeping." A previous study revealed that participants were able to correctly identify 25-ms bursts of periodic activity embedded within solar wind time series, and anecdotal evidence suggested that auditory scan may reveal events at much shorter time scales [192]. This is an excellent example of one such short-lived event occurring naturally within the STEREO data archive, and the fact that it was regularly identified through auditory scan supports the findings of the previous research.

One anomalous event that was identified by the expert in the first *Wind* data example was identified by all participants across both experimental conditions (this event is displayed in **Figure 35**). Several similarities can be drawn between this feature and the anomalous event displayed in **Figure 33**—both are book ended by regions of relative inactivity, and both display a general upward trend in frequency. The most striking difference between the two features lies in the relative absence of low-frequency spectral power in **Figure 35** (between approximately 100–300 Hz). Typical broadband solar wind

turbulence displays a power spectrum with a gradual decrease in amplitude/intensity from low to high frequencies, as energy emitted from the sun dissipates from large to small structures (i.e., large to small time scales). Participants' search schema likely incorporated this general behavior as *normal*, and in the case of the feature presented in **Figure 33**, this uneven spectral slope was flagged as *anomalous* through visual assessment. Close visual inspection of **Figure 33** reveals that this event also contains a slight enhancement in spectral power between 300–500 Hz, though in the presence of strong low-frequency turbulence this enhancement is not readily apparent to the eye. The ear, however, had relatively little trouble extracting this behavior as a separate auditory stream, as evidenced by the results presented here.

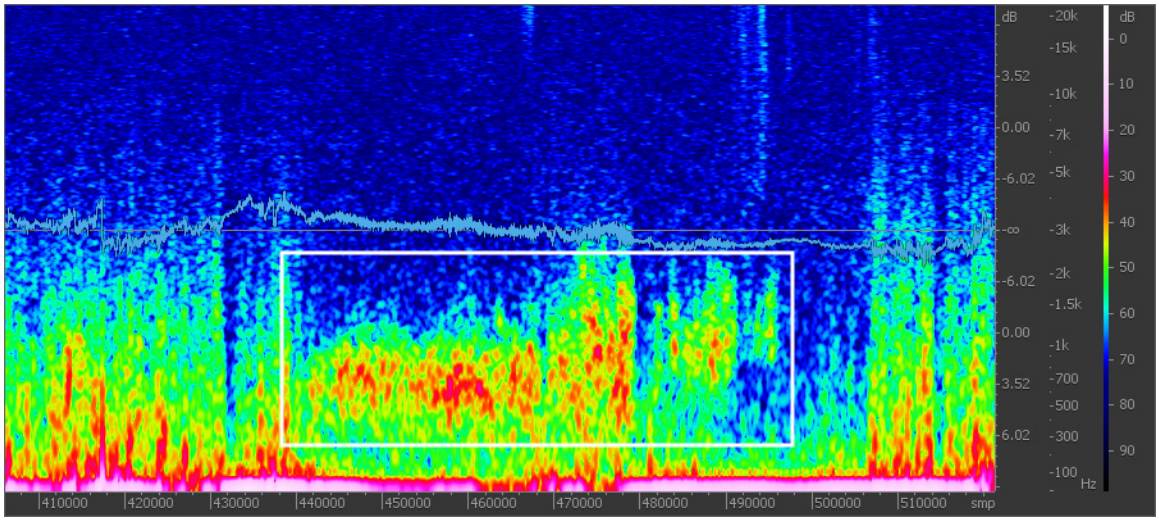


Figure 35. An anomalous feature in the first *Wind* data example (displayed in iZotope Rx) that was identified by all participants across both conditions.

3.3.8. **Analysis of descriptive vocabulary**

Participants were provided with a specific procedure for cataloging groups, anomalies within groups, and purely anomalous features, however, they were also encouraged to add a descriptive label for the purposes of clarification and recall. Several unique and highly-creative descriptors were provided for spectral features, such as "mountain" and "crunching snow." It was found that several labels were commonly utilized across all participants in the two search tasks—**Table 6** includes a list of the ten most common of these labels, presented in order from highest to lowest occurrence.

Table 6. The ten most common labels used to describe solar wind phenomena in the grouping task.

AUDITORY	VISUAL
Wind	Frequency
Rumble	High
High	Low
Sound	Spike
Scratch	Mid
Flag	Waveform
Noise	Very
Frequency	Dip
Background	Blob
Low	Shock

A quick scan of both lists reveals several terms common to both, including "high," "low," and "frequency." These terms would equally appropriate for a sound engineer describing the spectral content of an audio file or a space physicist describing features unfolding in solar wind time series. Perhaps more noticeable is the marked difference between the nature of selected terms, as the auditory list contains a variety of descriptive terminology and active verbs. Several participants likened the sound of the broadband solar wind turbulence to that of a "flag waving in the wind," while several anomalous events were reported as sounding characteristically "underwater." It should also be noted that several terms applied in the auditory description of solar wind (such as "rumble") could also be accurate descriptors for the sound of terrestrial wind.

Labels produced through auditory analysis tended to incorporate more experiential and emotional language—one particular region was labeled "sailing," and the participant noted that the sound reminded them of an experience on the ocean. Other regions were identified with colorful descriptions including "candle blowing out," "squeaky door," and "damaged tape." As participants had no prior exposure to this type of auditory stimuli they readily drew sound cues from memory—it is this same process that provided names for common heliospheric phenomena such as "lion roars" and "hiss."

3.3.9. **Detailed Analysis of the Grouping Task**

The results presented in this section do not take the expert classification into consideration; rather, they provide some insight as to the groups, within-group anomalies,

and purely anomalous events labeled by participants in the analysis task. Therefore, these data provide no measure of accuracy or agreement, and are reported here for the sake of better understanding of participant classification schema.

All but one participant created more groups in the visual task than the auditory task. Participants identified a total of 57 groups across all auditory examples, and 86 across all visual examples, for the identification of within-group anomalies these numbers dropped to 25 and 39 (respectively), and for pure anomalies these numbers dropped again to 22 and 11 (respectively).

Six participants identified more groups in the visual analysis task, one participant identified an equal number with both modalities and one identified more in the auditory task. Participants identified an average of 3.6 ± 0.7 groups per example in the auditory analysis task, and 5.4 ± 1.9 groups per example in the visual analysis task, a difference of 51% ($p = 0.05$).

Half the participants identified more within-group anomalies in the auditory task, and half identified more in the visual task. Participants found an average of 1.6 ± 1.4 within-group anomalies in the auditory analysis task and 2.4 ± 3.8 in the visual task, a difference of 56% ($p = 0.52$).

With regard to purely anomalous features, five participants identified more events through auditory analysis, two identified an equal number with both modalities, and one identified more events through visual analysis. Participants found an average of 1.4 ± 0.9 purely anomalous features per each auditory example, and an average of 0.7 ± 0.7 per each visual example ($p = 0.07$). Across both modalities, these results indicate that participants, on average, identified more groups than within-group anomalies. Participants also identified more within-group anomalies than purely anomalous events in all instances except in the auditory analysis of the first *Wind* data example. In every data example, visual analysis resulted in the identification of more groups than auditory analysis, while auditory analysis resulted in the identification of more purely anomalous events (when results are averaged across participants for each example).

3.3.10. **Assessing the Impact of the Audio Waveform**

In **Figures 33, 34, and 35** the blue audio waveform can be seen superimposed on top of the spectrogram display. For the auditory portion of the study, the spectrogram display was turned off, while the audio waveform remained on screen. In this way, the spectrogram visualization was directly compared against the auditory representation of the frequency spectrum, with the audio waveform serving as a common denominator in both conditions. However, it was observed that participants were not truly conducting a purely auditory analysis, as some visual information could also be extracted from the audio waveform display. A follow-up study was conducted in order to measure the impact of the audio waveform on the results of the auditory analysis task.

The experimental design of the original study was replicated with a fundamental change in the presentation of the stimuli. Eight participants were first provided the original visual and auditory grouping task, with one data example for each condition. For the second half of this study the audio waveform was completely removed and participants were again asked to assess one data example for each condition. The results across the two conditions (waveform present and waveform removed) were compared for any potential bias imparted by the presence of the waveform display; no statistically significant difference was observed in completion time, number of anomalous features identified, or overall agreement with expert assessment. **Figure 36** displays the average overall auditory agreement for both conditions across all four data examples. It was therefore concluded that the presence of the waveform display did not have a significant impact on the original anomalous feature identification task, and the results of the original study were presented as “auditory” analysis (rather than multimodal analysis) as it is believed that participants weighted their assessments primarily on the auditory feedback, and the waveform simply provided a visual reference for navigating the file.

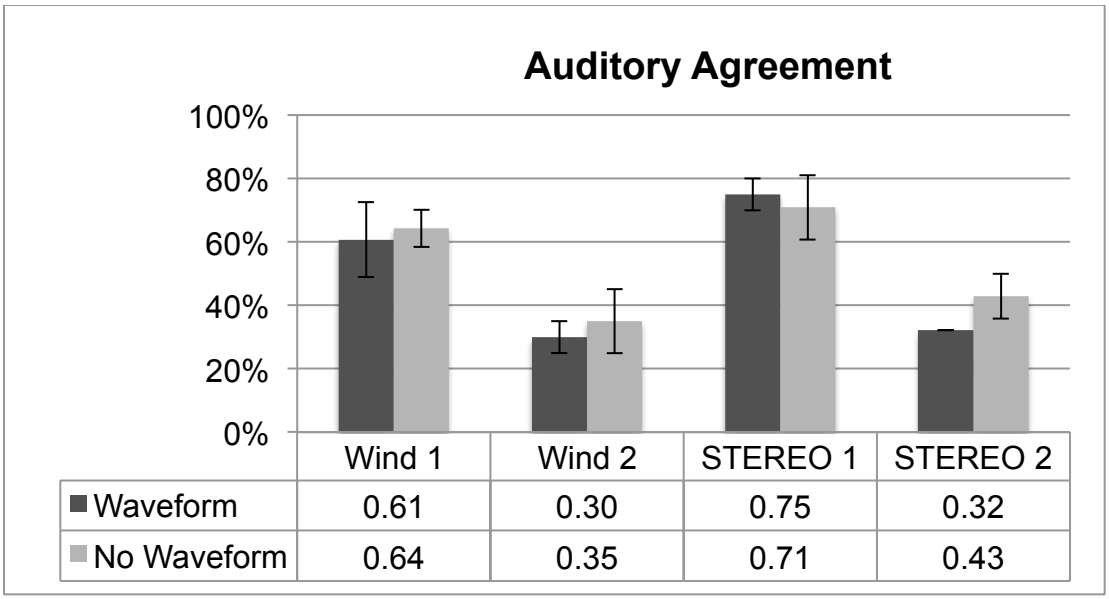


Figure 36. No significant change was found in participant agreement with expert assessment when the audio waveform was removed.

3.3.11. **Evaluating subjective feedback**

The post-test provided a free-response space for participants to comment on the comparative difficulty of the visual and auditory assessments—one participant noted that they found it "much easier to distinguish groups and anomalies with the auditory task." This feedback is unsurprising in light of the results that indicate a level of relative ease in the auditory identification task. Another participant found that they gradually got better at hearing differences in subsequent data examples—this points to the fact that most participants had no prior training or experience in auditory analysis and hence some amount of learning took place as participants learned to auditorily differentiate groups and anomalous features across several data examples.

One participant indicated that they found the listening task to be comparatively more difficult, and that auditory scan felt "very limited." A similar sentiment was expressed by several participants in previous studies presented in this chapter. There is no doubt that visual scan will remain the standard for some researchers who are more comfortable working with traditional analysis techniques.

3.3.12. **Limitations and Future Research**

Statistical significance was derived from the results of eight participants, a very

small sample size by the standards of the psychological sciences [203, 204]. The size of the participant pool was limited by the number of researchers in the heliospheric science division at NASA GSFC who regularly apply spectral analysis techniques in the evaluation of long time series. Each participant received a total of four data examples—two assessed through visual scan and two through auditory analysis. Consequently, this experimental design yielded four assessments per condition for each data example and a correspondingly low statistical power. The outcome of the study was also presented in absolute terms (participants using visual analysis overlooked 7 out of 8 anomalous features identified by the expert) such that the results may also anecdotally support the diagnostic power of auditory analysis techniques. A future study of this nature should incorporate a larger subject pool in order maximize statistical power and transferability.

This study was conducted within a proprietary software platform and hence the results concerning task completion time may only be considered from a relational standpoint across the two experimental conditions, as a similar spectral analysis task may take more or less time to complete using traditional scientific software platforms (e.g., Matlab or IDL). An intriguing future study should compare the rapidity of spectral analysis conducted within iZotope Rx to a similar analysis conducted within traditional heliospheric data analysis software—this would shed light on the relative efficiency of spectral analysis conducted with software tools that are specialized for the manipulation and exploration of audified data.

3.4. Conclusion

The work presented in this section demonstrated that participants using auditory analysis methods displayed a greater likelihood of detecting anomalous (and potentially novel) features in solar wind time series when compared to performing the same task using only visual feedback. This study incorporated data sets from ongoing missions and a participant pool consisting of active researchers, these two factors increase the generalizability of the results across the heliospheric science community, and it is hoped that these efforts may catalyze some degree of transferability across the scientific community at large.

The following two chapters depart from this quantitative investigation to address

the second research question, namely whether audification may reveal features that would otherwise be overlooked. Toward this end, domain-specific collaborative research is offered as anecdotal support for the hypothesis that auditory analysis can indeed reveal new features within complex time series.

CHAPTER IV

Audification As a Tool for the Spectral Analysis of Time Series Data

This chapter contains significant portions of R. L. Alexander, S. O'Modhrain, D. A. Roberts, J. A. Gilbert, and T. H. Zurbuchen, "The Bird's Ear View of Space Physics: Audification as a Tool for the Spectral Analysis of Time Series Data," *Journal of Geophysical Research: Space Physics,* 2014JA020025, 2014 [205]. This reserach is the primary qualitative contribution of this disseration. Several interactions with a single research scientist are documented as they gained the ability to extract relevant information from auditory cues. More than a set of anecdotal observations, this chapter provides new insight into audification as a working process in the domain sciences, and the ensuing discussion explores the impact of cross-modal cues and emergent descriptive language.

4.1. Introduction

This section has two primary objectives: **1)** to demonstrate the utility and a practical application of this relatively uncommon analysis technique, and **2)** to equip scientists with the necessary toolset to independently apply audification in the analysis of data gathered from a wide variety of sources (refer to **Figure 48** for a step-by-step guide). To achieve these two objectives, this chapter will briefly review the foundations of auditory analysis in the space sciences before presenting a case study in which a research scientist (henceforth referred to as 'the participant') was introduced to audification as a tool for the spectral analysis of one-dimensional time series data. In order to assess the impact of the auditory display, it was crucial to first understand the participant's traditional working practices. Toward this end, a preliminary workflow assessment uncovered a set of questions that the participant defined as guiding his approach during

the early stages of time series data analysis. These questions are generally focused on the integrity and scientific meaning of the data in question:

1) *Are the data well-formed and free of errors and gaps?*
2) *What large-scale structures are observable in the time series?*
3) *Can specific time regions that warrant in-depth investigation be identified based on a specific scientific rationale? (e.g., the presence of unique wavelet signatures, recurring structures, or outlying values.)*

The previous success of audification in the investigation of SWICS data suggested that it might support early diagnostic evaluations motivated by questions such as these. In order to evaluate the use of audification for the purposes of exploratory analysis we used a methodology that is commonly applied in the fields of interface design and usability testing, known as the "Think-Aloud protocol." This protocol requires a participant to vocalize their thoughts while engaging in a problem solving task, with the hypothesis that these verbalizations can reveal a subset of underlying cognitive processes [206]. Our goal was to assess to what extent the participant could use information gleaned from auditory observations to systematically drive the investigation of potential features of interest embedded within the time series. Towards this end the Think-Aloud protocol provided immediate insight as the participant's attention shifted from one auditory observation to the next. The nature and depth of these observations is the primary focus of this study, and two additional use cases are also presented. The findings support the hypothesis that auditory analysis can be useful in the identification of spectral features embedded in large time series data sets that may have been otherwise overlooked.

This work is conducted in collaboration with the Solar and Heliospheric Research Group at the University of Michigan and with a group of scientists at NASA Goddard Space Flight Center, currently the world's largest Heliophysics-focused research entity. The ultimate goal of this extensive research is to design both a methodology and a specific analysis tool through which space researchers can effectively apply audification techniques in the exploratory analysis of heliophysics-focused time series data.

It is important to note that, upon completing the initial study, the participant continued to use audification in the analysis of high-resolution magnetometer data from the *Wind* satellite on a weekly basis, and at present a collaborative investigation of a feature uncovered through auditory analysis is under review with the Astrophysical Journal [207]. This research therefore provides a detailed description of the initial steps towards a larger analysis framework through which data audification methods may be transferred successfully across scientific domains. A step-by-step guide for any scientist who is interested in applying this tool-set can be found in **Figure 48**, and a set of routines for audification in Matlab and IDL can be found online [208].

4.2. The Think-Aloud Protocol

To date, audification has been applied in the heliospheric sciences on an ad-hoc basis. This study is a formal investigation into how this technique may be regularly used by a researcher in the evaluation of large time series data. Accordingly, it was necessary to conduct a pilot study to observe the process of auditory analysis in action in order to gain an understanding of how it might guide a larger investigation. To accomplish this goal, a single case study employing the Think-Aloud protocol was undertaken to gain access to the reasoning process of the participant in real-time [209]. Originally introduced by Lewis Clayton at IBM in 1982 [210], this verbal response protocol has remained a cornerstone of software engineering research, and successful use cases can be found across all stages of the design process [211]. Notably, the Think-Aloud protocol has been used to assess the effectiveness of various forms of knowledge representation and visualization [212] and to evaluate the usability of new software tools during early prototyping stages [213].

4.3. Origins of the Case Study

Sonification techniques have proven particularly useful for the exploratory analysis of various types of scientific data [214]. Exploratory data analysis is an open-ended process that involves making large datasets more easily navigable to the human analysts, with the goal of gaining knowledge and uncovering new insight. As opposed to confirmatory data analysis, which seeks to assess how well our assumptions align with available data, exploratory analysis focuses on the acquisition of knowledge that lies

outside our realm of expectations [9]. The current study is a portion of a more extensive project to evaluate the application of audification for exploratory data analysis of time series. Previous research indicated that audification has the potential to reveal features that may elude the eye [16]; this chapter documents multiple cases that support this hypothesis in order to provide a better understanding of the unique spectral characteristics of features that are predisposed to auditory identification in these instances.

4.3.1. **Data Selection**

The team worked with the participant to determine a suitable heliospheric data set for exploration through auditory analysis. Generally speaking, data sets appropriate for audification include large one-dimensional time series with samples gathered at regular temporal intervals. Magnetometer data from the *Wind* spacecraft presented itself as an ideal candidate based on its high (92 ms) time resolution, the variety of spectral features, and the length of continuous data, which extend longer than one solar cycle. This dataset enables scientific investigations focused on the microphysics of the solar wind, e.g., the interactions between waves and particles in the turbulent solar wind—fundamental processes that indicate how energy in plasmas is transported and dissipated [215, 216]. Certain spectral features in the data can indicate the presence of specific types of waves, such as cyclotron waves, which are a proxy for strong wave-particle interactions [217]. Such wave-bursts are often short lived and difficult to find through traditional analysis methods, particularly considering the large volume of available high-resolution data. Audification provides a promising alternative, under the hypothesis that these spectral features may be located by listening to a large quantity of data in audio-file format.

4.3.2. **Data Cleaning**

The first diagnostic evaluation presented in the introduction related to the assessment of errors and gaps in the time series. Listening through several examples of audified magnetometer data quickly revealed that these values can often produce audible “clicks” and “pops.” An initial step in the audification process was to determine the best way to manage these “bad” and missing data values. It was found that in instances where a small percentage of data are missing, linear interpolation across these values may create minimal auditory artifacts while preserving the time scale of the original data set. An

alternative approach would be to assign these entries a pre-determined value (such as zero), or simply remove them from the audified data altogether [218].

4.3.3. **Preliminary Analysis**

The team began with a data survey of *Wind* high-resolution magnetometer data from 2007 and 2008. These years were prime candidates for the presence of wave activity due to the relative period of inactivity during solar minimum. During the early stages of the study, we used auditory analysis to identify potential wave activity, and a set of specific time regions were provided to the participant for evaluation through traditional analysis methods. Upon repeated exposure to many audified data examples, the participant developed the ability to consistently identify wave activity through the application of analytical listening techniques. The following section explores the nature of this ability in a structured analysis task.

4.4. Structured Think-Aloud Study

A Structured Think-Aloud is employed to gain access to the evolving reasoning processes of the participant during a multimodal analysis task. This protocol includes the use of predetermined verbal prompts that are provided during an extended period of inactivity (e.g., 5 sec) [219]. The central hypothesis of this study is that through analytical listening, a research scientist will be able to successfully identify wave activity within audified magnetometer data sets. The participant is a 30-year-old male with a physics Ph.D. and extensive experience working with spectrogram displays. He has no self-reported hearing or vision impairment. Consent was given to record audio during the Think-Aloud task, and the University of Michigan deemed the study to be exempt from the oversight of the Institutional Review Board (IRB).

High-resolution *Wind* magnetometer (MAG) data gathered during November 2007 were downloaded from the CDAWeb data repository (CDAWeb, Goddard Space Flight Center; http://cdaweb.gsfc.nasa.gov/cdaweb/sp_phys/). These files were combined in Matlab and audified as a single 16-bit file in *.wav* format, and two time regions were selected for analysis. All audio examples discussed in this chapter can be found online [208].

All listening tasks were completed with Audio Technica ATH-M50 professional studio monitor headphones. The auditory analysis task was completed on a Macbook Pro using the iZotope Rx 2 software platform. All instructions were verbally provided to the participant.

The participant was asked to think aloud while engaging in a directed multimodal data analysis task. Presentation of the data began with playback of the audio file, and the participant was prompted with questions such as "Can you describe what you're hearing in words?" and "Would you like me to replay any certain section?" This analysis was largely exploratory in nature, as the participant was not specifically asked to identify certain types of features. This provided important information as to the type of event that the participant would auditorily identify as a feature of interest. Once the initial auditory analysis was complete, the participant was provided with the waveform of the audified data for visual reference. At this time, additional prompts were provided such as "Is there anything interesting about what you're observing?" and "Would you like me to zoom-in, zoom-out, or move to a different region?" Finally, the spectrogram display was made available, and the participant was asked to continue verbalizing his thoughts as he used the visual and auditory displays simultaneously. The study was conducted in a quiet room over the course of a single session.

4.5. Results

An audio recording of the Think-Aloud session was transcribed to produce verbal data. During the session, all instances in which the participant directly referred to the audified dataset were documented, and the corresponding data sample numbers were logged in the transcript. This allowed vocalizations to be paired with specific features in the data. Any instance in which the participant verbally reproduced the spectrum of the audified data was highlighted, along with any terminology used to describe the auditory phenomena. It was determined that segmentation of the entire verbal protocol was not necessary, as an initial search revealed two specific cases in which the participant anticipated the presence of wave activity based on auditory observations, and subsequently confirmed this activity through visual analysis. These portions of the experimental session have been selected for presentation in this section.

The data used in both instances were taken from the *z* component (in the GSE coordinate system) of *Wind*/MAG observations, as the *x* and *y* components contained an audible tone induced by the spinning of the spacecraft (while the *z* component did not). Both examples were measured in close proximity to the passing of a large magnetic cloud. The first example spanned 123,116 data samples from solar wind magnetic field measurements gathered on the 18th of November 2007. The resulting audio file was approximately 2.8 seconds in length at playback rate of 44,100 samples per second and represented approximately 3.1 hours of real-time recording.

When listening to the audified data, the participant was able to identify several instances of wave activity, and he expected that the spectrogram display would contain "…a low power, some kind of gradient with a peak at some specific range of frequencies for the chirp at the beginning." He then speculated that the middle portion of this example would contain one or more "peaks" at a lower frequency, and that they would "drift or there will be some change." Additionally, he indicated that the middle portion contained an event that was potentially "percussive" in nature, and marked by a sudden rise in amplitude. He anticipated that the latter half of the event would have a steep slope and the spectrogram should "lift up." In the previous description, a "peak" refers to a region with increased spectral power, and the "chirp" corresponds to the 1 Hz enhancement that can be seen near the left side of **Figure 37**. Here, a spectrogram of the audified data in iZotope Rx (top) is provided for direct comparison with a spectrogram of the original time series rendered in Matlab (bottom). Note that while the audio file has been transposed into the frequency range of human hearing, the two plots are nearly identical. In both representations, the "rise in amplitude" described by the participant can be seen as an increase in spectral power occurring in the latter half of the time interval.

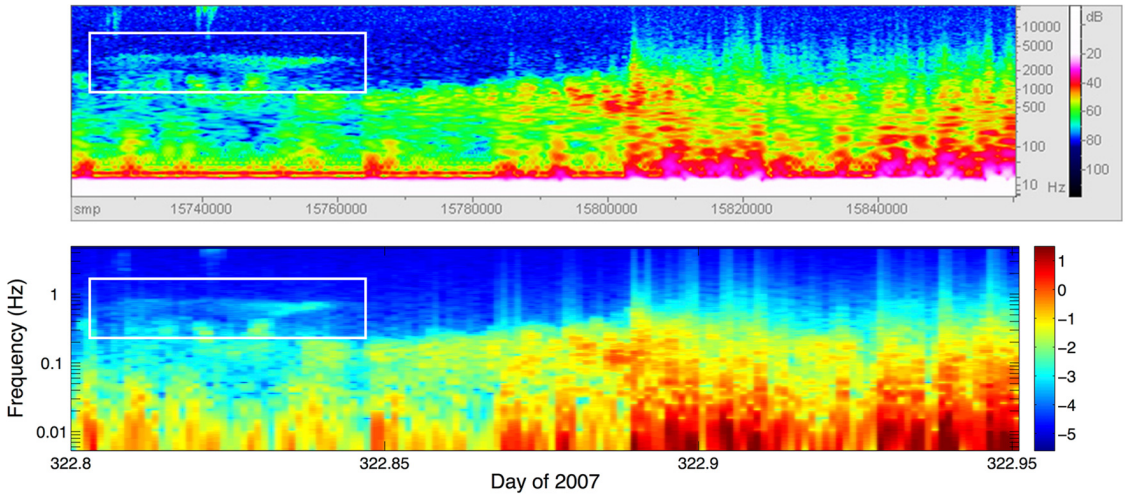


Figure 37. The spectrogram of the first data example rendered in iZotope Rx (top) and Matlab (bottom). This interval spans 123,116 data samples from solar wind magnetic field measurements gathered by the *Wind* spacecraft on 18 November 2007 (DOY 322). Here the participant described a "chirp" event corresponding with the band of 1 Hz activity near the left-hand side (see white box) (reprinted from Alexander et al. 2014, Figure 2, p. 6).

The participant indicated that his previous experience working with audified magnetometer data informed his assessment of features in this example. Before conducting a visual inspection, he hypothesized that the "chirp" event would be associated with a region of coherent wave activity. The original hypothesis was confirmed upon the observation of moments of coherent wave activity in the audified data waveform. The time series for this region is provided in **Figure 38**.

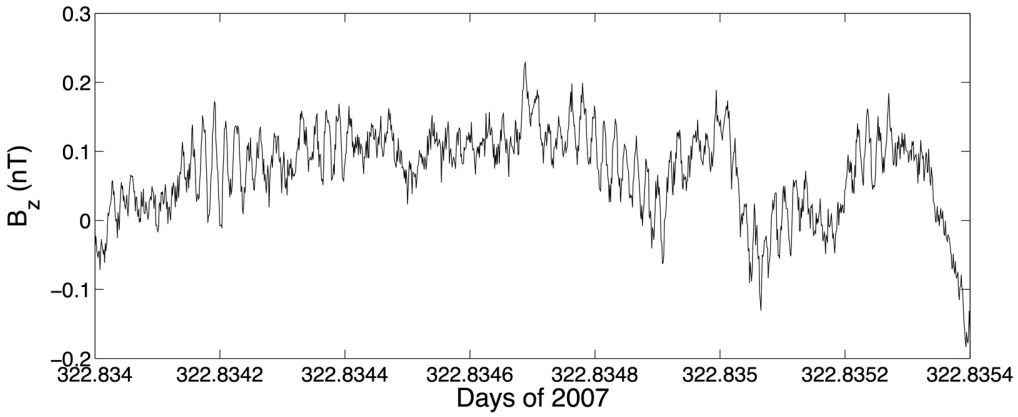


Figure 38. A sub-region of the event occurring in *Wind* magnetometer data that was identified as a "chirp" through auditory analysis. Instances of coherent wave activity can be seen in the time series as nearly sinusoidal oscillations (reprinted from Alexander et al. 2014, Figure 3, p. 6).

The second example spanned 148,837 data samples from magnetic field measurements gathered on the 20th of November 2007. The resulting audio file was approximately 3.4 seconds in length at the same 44.1 kHz sampling rate, representing approximately 3.8 hours of real-time recording. Upon listening to the audified data (without access to the spectrogram display) the participant observed three distinct sections: A "*warble*" noise leading up to a short "*knock*" at a slightly higher frequency, and finally a quieter segment containing broadband noise that was both rising and "*hissing*." He hypothesized that wave activity would present itself in small packets across a range of frequencies. The time series and spectrogram representation for the second audified data example are provided in **Figure 39**.

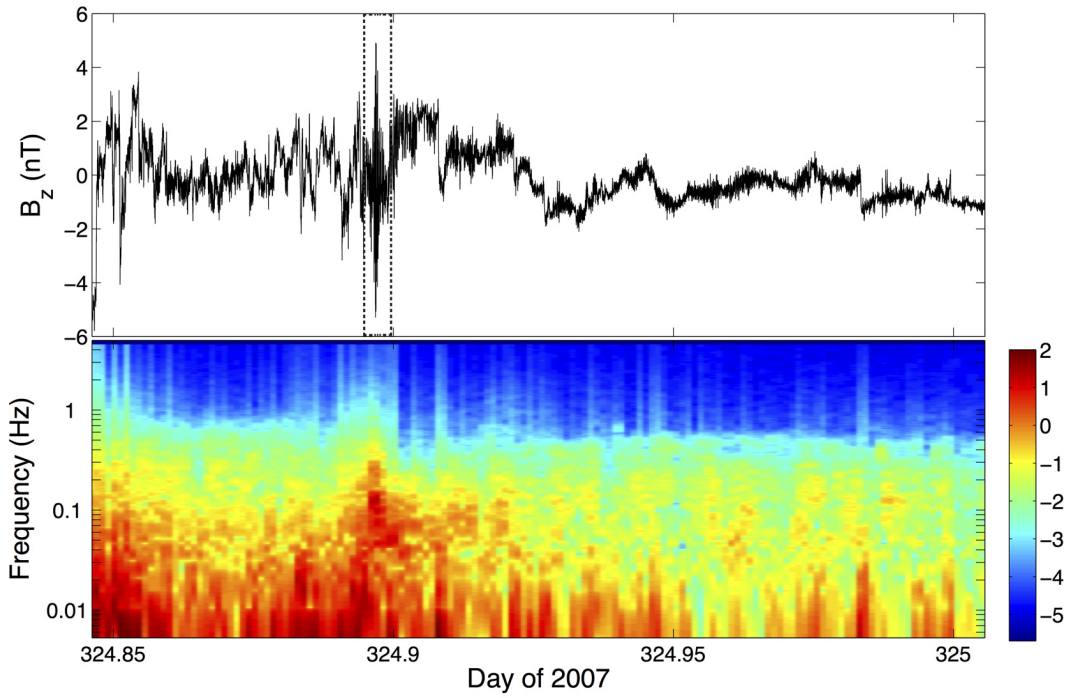


Figure 39. The time series (top) and spectrogram display (bottom) for the second audified data example. This interval spans 148,837 data samples from solar wind magnetic field measurements gathered by the *Wind* spacecraft on 20 November 2007 (DOY 324). The participant divided this example into three sections he described as a "warble" noise, a "knock," and finally a "hissing." A dotted line has been placed around the "knock" event in the time series, and this region is expanded in Figure 5 (reprinted from Alexander et al. 2014, Figure 4, p. 7).

The event described as a "knock" can be seen in both the time series and spectral display as a short increase in power occurring roughly at day 324.9. When subsequently

provided access to the audio waveform, the participant inspected this region using a combination of auditory and visual analysis methods to isolate the "knock" feature. Visual inspection revealed the presence of six relatively clear oscillations within a larger amplitude envelope. The time series for this region is provided in **Figure 40**.

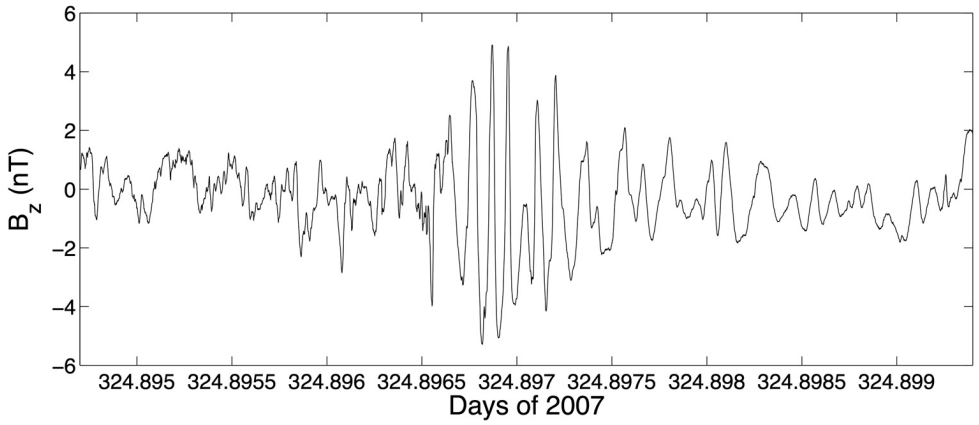


Figure 40. A subregion of the event occurring in *Wind* magnetometer data (z-component) that was auditorily identified by the participant as a "knock." Close inspection of the time series reveals six periodic oscillations within a larger amplitude envelope (reprinted from Alexander et al. 2014, Figure 5, p. 7).

Additional investigation of these two time intervals through traditional analysis methods confirmed the presence of wave activity. The first example likely contained waves caused by a stream interaction region and the second example contained the reverse magnetosonic shock of an interplanetary coronal mass ejection [220].

4.6. Discussion

In this case study, the research scientist was able to extract important spectral cues through the application of auditory and multimodal analysis methods. While a correlation between assessments made through audition and was established by Pauletto and Hunt [19], this case study further demonstrates how auditory analysis may be useful when applied in an open-ended feature identification task.

4.6.1. **Cross-Modal Cues**

Across the two data examples the participant related instances in which features were easier to detect through either visual or auditory analysis. In the first data example,

some portions of the frequency spectrum only became audible upon viewing the spectral representation. Specifically, the participant stated that once he could see the spectrum he was able to hear a lower frequency feature that had eluded identification during initial auditory analysis. He indicated that he had been distracted by a simultaneous "cheep" event that occurred at a higher frequency. Here, the audified low frequency content remained peripheral until it was pulled into focal awareness by visual observation. This underscores one of the potential strengths of multimodal analysis: one sense may cue the other into the presence of important information that may otherwise be lost.

In the analysis of the second example, auditory observations prompted the participant to conduct additional visual exploration of the spectrogram display. In this instance, the participant indicated that he could hear a feature that was imperceptible in the spectrogram, and through a subsequent re-scaling of the data he was able to visually confirm the presence of a subtle narrow-band spectral enhancement.

During the Think-Aloud session the participant noted several examples of wave activity that were visible in the audio waveform (the presence of waves was later confirmed through traditional analysis methods). This speaks to the isomorphic nature of the audification process—as all data samples were preserved, the participant readily regarded the audio waveform as a one-dimensional line-plot of the original data (see **Figure 1**). In several instances, upon detecting a feature in the waveform the participant immediately associated the region with a specific sound that was observed during audio playback.

4.6.2. **Descriptive Language and Vocalization**

Another way to understand how the participant categorized features he heard within the data is to observe the language he used to describe these features and, in some cases, how he attempted to imitate the sounds he heard. Throughout the experimental session, the participant used a variety of descriptive techniques in communicating the form and structure of features within the audified data. Wave activity was described in terms including *warbles, whooshing, swirling, chirping*, and *whirling*. In several instances the audified data were related to familiar acoustic phenomena such as a knock, a spinning hollow tube, or sounds derived from a metallic cable. On more than one

occasion, nearby objects were used to convey acoustic information: in the instance of the "knock" event, the participant slapped the top of the desk to indicate the percussive nature of the sound.

Another strategy involved using the mouth as a complex filter to vocalize the nature of the evolving spectrum. In these instances hard consonant sounds such as "ck" and "ch" relayed the presence of transient broadband noise. The presence of narrowband noise was conveyed with exhalation while opening and closing the mouth, indicating the presence of low- and high-frequency content respectively.

4.6.3. **General Discussion**

In the month prior to the Think-Aloud study, the participant was exposed to several examples of audified *Wind*-MAG data. He noted that he had developed the ability to auditorily detect spectral components that occurred close to the cyclotron frequency, and that these specific sounds would lead him to suspect that wave activity would be present in the data. This pre-exposure was important in developing the ability to auditorily distinguish "normal" solar wind turbulence from abnormal behavior (in many instances a new vocabulary was necessary to describe abnormal activity that had yet to be defined in scientific terms). This study demonstrates that the participant's listening abilities were sufficiently developed to associate specific auditory observations with certain types of wave-particle interactions, supporting the earlier suggestion that audification can assist in the identification of time regions that contain features of interest. Though the participant had never been exposed to the auditory examples provided in the Think-Aloud, the spectral content was sufficiently familiar to classify wave activity based on the participant's acquired "auditory vocabulary." The ability to reliably connect auditory observations with precise regions of the original dataset relies on the fact that certain features in the solar wind (e.g., ion cyclotron waves) will consistently give rise to similar spectral features in audified data, and the resulting auditory streams, which Bregman [65] refers to as auditory objects, are closely tied to the evolving physical phenomena.

Another key question posed in the introduction was whether audification could support the observation of large-scale structures in the time series. The level of detail

provided by the participant when asked what he expected to see in the spectrogram display indicates that he was able to both auditorily extract specific features and also observe the wider frequency spectrum as it evolved. This type of auditory scanning may be particularly well suited for the purposes of exploratory analysis, as potential features of interest may be present across a wide frequency range. Furthermore, as the initial assessment of the frequency spectrum was conducted without any visual reference it is feasible that the eyes could be engaged in a secondary task, though additional research is necessary in order to quantitatively assess the impact of multimodal displays when applied toward an exploratory analysis task.

4.7. Two Additional Examples

4.7.1. **Detection of Equipment-Induced Noise**

The participant independently audified 200 days of solar magnetic field observations gathered by the magnetometer on the Ulysses spacecraft, the resulting audio file has been uploaded to a web-based repository [221]. Through auditory analysis, he was able to detect equipment-induced noise that he had not observed previously. Specifically, he observed aliasing in the high-frequency range of the audified data, which was most likely introduced by the tape recorder on the spacecraft. As the two spools run at different rates, aliased "drifting tones" can appear in the data. (T. Horbury, personal communication, February 11th, 2014)

The participant stated that this noise was "something that stood out by ear more than it stood out by eye." This supports the suggestion posited by Hayward [1994] that one of the most promising applications for audification may lie in the identification of noise and equipment-induced error. The aliasing effects in the audified Ulysses data are observable as thin green lines sweeping across the high frequency ranges in **Figure 41.**

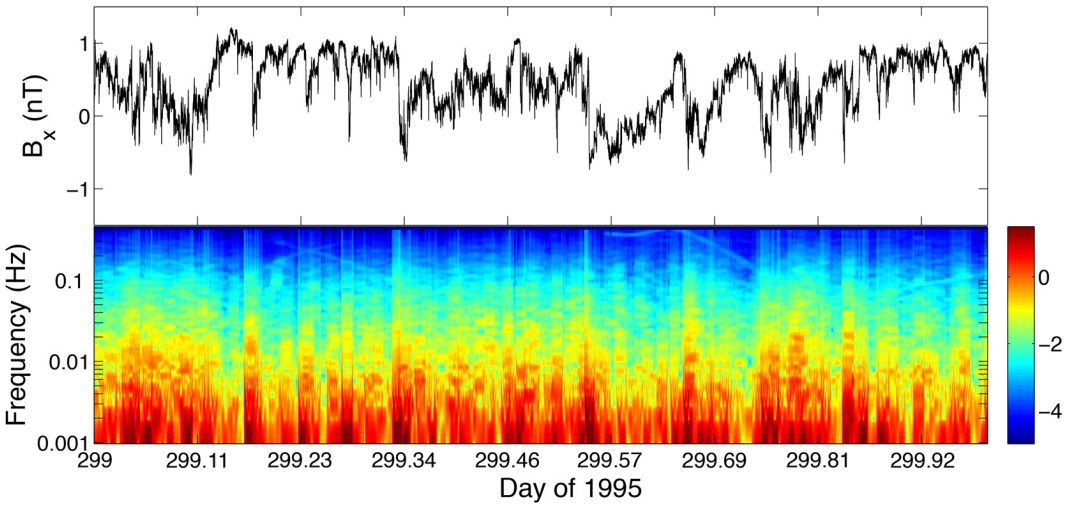


Figure 41. The time series (top) and spectrogram display (bottom) of data from the Ulysses magnetometer instrument. This interval spans 86,401 data samples from solar wind magnetic field measurements gathered on 26 October 1995 (DOY 299). The spectrogram contains aliased artifacts at high frequencies. These artifacts were first identified by the participant through auditory analysis, and later confirmed through visual analysis (reprinted from Alexander et al. 2014, Figure 6, p. 9).

4.7.2. **Manipulating the Audified Data**

In a subsequent analysis session, the participant hypothesized that the second example from the Think-Aloud study contained an instance of a reverse shock. A reverse shock occurs when a fast plasma stream is followed by a slower one, resulting in a shock wave that travels towards the sun within the reference frame of the plasma (while still traveling away from the sun in the reference frame of the spacecraft). If this event were indeed a reverse shock, the physical structures would be nearly identical to a forward shock but would evolve in reverse temporal order. Here the participant suggested that the audio for this reverse shock event should be played backwards, sped up, and compared with a forward shock to assess for similarity.

The audified data were subsequently processed in Audacity, a free software platform for editing and manipulating audio waveforms. After reversing the file, the team used auditory analysis in tandem with calculations of the anticipated Doppler shift to match the spectral contour with that of a known forward shock. The resulting audio was a convincing match, and the original hypothesis was confirmed. This audio example can be found online [208].

Many Digital Signal Processing (DSP) techniques used in the processing of heliospheric time series data have correlates in the digital audio domain. When carefully applied, DSP algorithms for noise reduction and filtering may potentially reveal new features within an audified data set. For example, it has been found that artificially induced frequencies from spacecraft spin-tone can often be auditorily attenuated through the application of a notch filter; and adjusting the sampling rate or playback speed of the audio will shift the entire frequency spectrum, revealing new micro and macro features inherent within the data.

4.8. Conclusion

In the case study reported in this chapter, auditory analysis techniques were successfully applied by a research scientist in the detection of wave activity embedded in 11 Hz *Wind*-MAG data. Though the participant had some familiarity with data of this type, the fact that he was able to auditorily recognize wave activity in regions of the data to which he had no previous exposure suggests that he had acquired a sort of "auditory vocabulary" for specific types of wave-particle interaction. Additionally, the participant was able to use audification to identify equipment-induced noise that had been previously unobserved. This research, therefore, suggests that audification can reveal spectral features in solar wind time series that can inform early diagnostic evaluations. More specifically, this study indicates that audification can be useful in the initial stages of overview and feature identification, providing a new "big picture awareness" that was not present before. This global perspective could be described as a *bird's ear view*, a macro understanding of large time series brought about through auditory scanning, a process that preserves the small time scale features that may be difficult to see when performing a large-scale visual scan.

A follow-up interview with the participant indicated that he has continued to use audification in the assessment of every large high-resolution data set he has worked with in the two months since the study was completed, and that in many instances auditory scanning has been preferable to visual scanning methods. Further analysis of cases like those found here has led to the discovery of intervals of wave activity that have clear

correlations with other phenomena such as particle beams; these observations represent scientific advances that will be discussed in a future publication.

It should be noted that the observations in this chapter reflect an interaction with one research scientist and were motivated by a number of previous scientific studies in which audification led to important results when applied by a trained specialist (e.g., Landi et al., 2012). A generalized study that focuses on the broad applicability of audification as a data-analysis tool is currently in progress, and this ongoing investigation includes a significantly larger participant pool. We will analyze research scientists' visual and auditory observations in depth to better understand the specific strengths and weaknesses of each modality (as well as their interplay). This follow-up experiment has been designed to produce quantitative results that should be more readily transferrable to other scientific domains. Additional qualitative research will continue to assess the evolving workflow of research scientists as they begin to integrate auditory analysis tools and methods. Ultimately this work will inform the design of a tool for the multimodal analysis of large one-dimensional time series and the creation of an interactive web-based tutorial series.

While visualization has long been the standard technique for representing scientific data sets, this research indicates that audification can be a valuable diagnostic tool in the analysis of large time series data sets. This investigation is an initial step towards a larger framework through which data audification as a method for spectral analysis may be transferred across scientific domains.

The data sets used in this chapter are accessible through the CDAWeb data repository (maintained by NASA Goddard Space Flight Center); all audification examples in this chapter can be accessed online through the University of Michigan's permanent data archive [208].

CHAPTER V

Applied Auditory Analysis

Introduction

This chapter contains four case studies in which an audification specialist worked alongside members of the University of Michigan Solar And Heliospheric Research Group (SHRG) and research scientists at NASA Goddard Space Flight Center (GSFC). [28] In each case, audification was successfully applied to identify previously undiscovered features in high-resolution solar wind time series. These well-documented successful use-cases should provide a valuable reference within the growing body of sonification literature. The work reported in section one resulted in a publication with the Astrophysical Journal (ApJ), section two documents an auditory observation that resulted in both an accepted abstract with the American Geophysical Union (AGU) and a paper in progress, section three contains excerpts from a paper under review with ApJ, and section four contains collaborative research that was presented to the Committee on Space Research (COSPAR 2014). Key information regarding the nature of these four case studies is summarized in **Table 7**. It should be noted that audification was effectively applied **1)** across a wide range of data sources including observations of particles, magnetic, and electric fields, **2)** for both exploratory and confirmatory analysis, **3)** for data sets ranging from 1 to 30 dimensions, and **4)** for the evaluation of time series that varied in length from 1.5–344 million data samples. Therefore, this chapter demonstrates the potentially broad applicability of this analysis technique.

[28] For the remainder of this chapter, every use of the term "specialist" will explicitly refer to the audification specialist, while "scientist" and "researcher" will explicitly refer to domain scientists.

Table 7. A summary of the four case studies provided in this chapter.

Publication	Title	Data Source	Analysis Type	Data Dimensions	Data Samples
Landi et al. (2012) [222]	Carbon Ionization Stages as a Diagnostic of the Solar Wind	**ACE/SWICS** *Particles*	Exploratory	30	1.5 Million
Tang et al. (2014) [223]	First Simultaneous Observations of Lower Hybrid, Whistler-Mode, Electrostatic Solitary, and Electrostatic Electron Cyclotron Waves near the Earth's Magnetopause	**THEMIS/EFI** *Electric Fields*	Exploratory	3	2.3 Million
Wicks et al. (2015) [207]	A proton cyclotron wave storm generated by unstable proton distribution functions	**WIND/MFI** *Magnetic Fields*	Exploratory/ Confirmatory	1	334 Million
Jian et al. (2014) [224]	Magnetic waves near the proton cyclotron frequency in the solar wind	**WIND/MFI** *Magnetic Fields*	Confirmatory	1	344 Million

In each case study the investigation is framed within a larger domain-specific context and presented in stages of data selection and preparation, audification, auditory analysis, and knowledge extraction. This work is presented in chronological order, spanning approximately three years; over this period the nature of the analyses shifted from exploratory (case studies 1 and 2) to confirmatory in nature (case studies 3 and 4). The large number of new domain-specific research initiatives stemming from this work suggests that the investigation of large time series data archives benefits from multimodal data mining. While audification is demonstrated to expedite the feature extraction process in confirmatory analysis (case studies 3 and 4), it is more surprising that significant value was generated through open-ended exploratory search (case studies 1 and 2).

5.1. Carbon Ionization Stages as a Diagnostic of Solar Wind Source Regions

Section 5.1.4 contains an auditory analysis originally published in R. L. Alexander, J. A. Gilbert, E. Landi, M. Simoni, T. H. Zurbuchen, and D. A. Roberts, "Audification as a Diagnostic Tool for Exploratory Heliospheric Data Analysis," in *17th International Conference on Auditory Display (ICAD 2011)*, Budapest, Hungary, 2011 [225]. Sections 5.1.4 and 5.1.5 contain excerpts from E. Landi, R. L. Alexander, J. R. Gruesbeck, J. A. Gilbert, S. T. Lepri, W. B. Manchester, and T. H. Zurbuchen (2012), Carbon Ionization Stages as a Diagnostic of the Solar Wind, *The Astrophysical Journal,* volume 744, p. 100 [222]. This publication resulted from an interaction between an audification specialist and members of the Solar and Heliospheric Research Group at the University of Michigan.

In an initial meeting with the SHRG the specialist was provided with general information regarding basic heliophysics (e.g., the distinction between slow and fast wind, and between different solar structures such as sunspots, coronal holes and active regions) and directed to download data generated by the Advanced Composition Explorer (ACE) spacecraft from a publicly accessible data repository [226]; over 20 parameters were subsequently evaluated for subtle spectral cues. A prominent underlying hum was identified, and the audification specialist was able to attribute this hum to the solar synodic rotation, confirming the rotational period with a margin of error below 1 percent. One data parameter was offered to the research group as a prime candidate for additional analysis as it displayed an unusually strong harmonic signature that had yet to be formally examined. Collaborative investigation by the research group revealed that this data parameter provided a powerful tool for identifying the type and source-region of the solar wind arriving at ACE—this parameter (the C^{6+}/C^{4+} ratio) was more powerful than the tool that had been routinely used for decades (the O^{7+}/O^{6+} ratio). The following case study is offered as an example in which auditory observations were the catalyst for a new line of research within the domain sciences.

5.1.1. **Scientific Rationale**

The SHRG is the world-leader in the analysis of solar wind elemental and ionic charge-state composition. Datasets produced by sun-observing satellites provide a wealth of new knowledge regarding the origin and acceleration of the solar wind, but can be difficult to navigate by non-experts [227]. Early interactions between the audification specialist and the SHRG were driven less by a domain-specific hypothesis, and more by the broad question of whether sonification could successfully reveal new features in the vast archive of heliospheric time series data. Toward this end, the audification specialist began an in-field immersion in order to gain domain-specific knowledge and vocabulary. Weekly meetings with the SHRG were attended at which the specialist was encouraged to actively participate and ask questions. The specialist also completed a graduate-level course in the department of Atmospheric, Oceanic, and Space Sciences specifically focused on the evolution of physical processes in the Sun and the Heliosphere. Additional one-on-one meetings with research scientists assisted in establishing a fundamental knowledge base. *The specialist published the following brief introduction to solar wind types and ionic charge state compositions in the proceedings of the 17th International Conference on Auditory Display [225].*

The ability to forecast space weather has extensive benefits, as accurate predictions of solar storms can protect astronauts from health hazards due to extreme space environment conditions. Accurate forecasting also allows preventative measures to be taken towards minimizing damage to delicate instruments, as solar storms can disrupt satellites and interfere with ground communication.

The occurrence of solar storms directly correlates with solar activity, which oscillates in an 11-year cycle. We can glean insight into this cycle by closely studying the solar wind. With tools such as the Solar Wind Ion Composition Spectrometer, we are able to analyze the composition and determine the source of solar wind plasma. Coronal hole wind, also known as “fast” wind, has an average bulk speed of 750 (km s^{-1}). This wind originates from coronal hole regions, which are areas of low temperature (~0.8MK) and open magnetic flux located primarily at the poles during solar minimum (and at lower latitudes during solar maximum). Non-coronal hole wind, also known as “slow” wind,

has an average bulk speed closer to 400 (km s^{-1}) and comes from hotter regions (~1.3MK) in the solar corona.

These two types of wind are more accurately identified by their charge state composition than their speed [228, 229]. In fact, temperatures derived from observed charge states are unique for different elements [230] and indicate the temperature of the wind at the height where charge states "freeze in." (**Figure 42**) They provide invaluable information as to the solar wind type, source region, and acceleration mechanism. Currently, the O^{7+}/O^{6+} ionic charge state ratio is utilized as a tool to distinguish between differing types of solar wind plasma based on their freeze-in temperature [231].

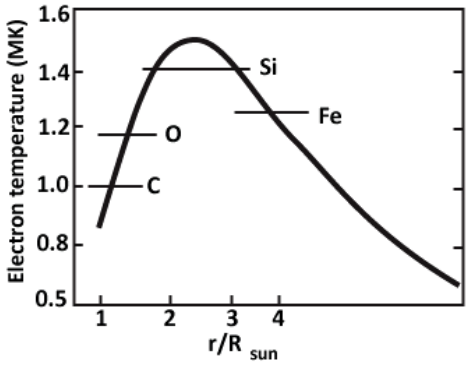


Figure 42. Electron temperature, measured from charge states, as a function of distance from the sun. Figure adapted from Geiss et al. (1995) (Reprinted from Alexander et al. 2011).

5.1.2. **Data Selection and Preparation**

The work of the SHRG is primarily focused on observations gathered by the SWICS instrument on the ACE satellite. This instrument—constructed in part by members of the SHRG—is optimized for measuring the chemical and charge state composition of solar (and interstellar) particles and has produced a growing data archive since 1997 [227]. This archive is provided on a publicly accessible repository actively maintained by the ACE Science Center at the California Institute of Technology; available data parameters include density, bulk velocity, thermal velocity, and ratios of charge state densities and elemental abundances for over 60 ions [232]. These data are available in time series format with a sampling rate of 1-hour, 2-hours, or full daily averages. A flexible web-based interface allows for the retrieval of various solar wind

data parameters with customizable time-formats. Bad and/or missing data are provided an error flag and reassigned to a value of "-1."

5.1.3. **Audification**

Yearly data files containing solar wind density, velocity, and charge state values at a 2-hour sampling rate were downloaded from the ACE/SWICS repository and combined to create a single file spanning 1998 to 2009 inclusive. This file was imported into the "text" object in the Max/MSP computer-music programming environment and the "route" object was used to filter any bad or missing data. Each parameter was scanned for maximum and minimum values, which were used for scaling the data between -1 and 1 for writing directly to an audio buffer. This process generated more than 20 single-channel 16-bit audio files in .wav file format. The following analysis explores the investigation of audio files produced from C^{4+}, C^{5+}, and C^{6+} charge states, parameters that the specialist noted were rarely discussed during meetings with the SHRG.

5.1.4. **Auditory Analysis**

While these audio files initially appeared to be quite noisy, close listening revealed an underlying "hum" with a frequency of 137.5 Hz during times near the solar minimum of solar cycle 24. We compared this hum with the sound of a sinusoidal oscillator of adjustable frequency, and found agreement for a frequency of 137.5 ± 1.0 Hz. Such a frequency translates to a 26.94 ± 0.20 day period within the original data, which corresponds to the synodic rotation period of the Sun, in excellent agreement with the mean equatorial synodic rotation rate estimated by Newton & Nunn (1951) using recurring sunspots [233].

Close listening to these charge states also revealed that their spectral characteristics shared a unique "flavor" that was difficult to discern through visual analysis. Strong partials above the fundamental frequency (137.5 Hz) were heard at 275 Hz and 550 Hz. With a known fundamental periodicity of 26.4 days (in the original data), the first three harmonics were calculated as occurring with periodicities of roughly 13.2, 8.8, and 6.6 days. These periodicities were noticeably stronger in the audification of C^{6+}

and C^{4+}, and less present in C^{5+}. This periodicity was related to a triad of equally spaced coronal holes on the surface of the Sun [234].

5.1.5. **Knowledge Extraction**

Figure 43 shows the Fourier power spectrum of three quantities routinely measured by ACE/SWICS as a function of time in the solar wind: solar wind velocity (measured as the bulk speed of He^{2+}), the C^{6+}/C^{4+} ratio, and the O^{7+}/O^{6+} ratio. The He^{2+} velocity, measured by SWICS, was compared with the H, C, and O velocities over the entire ACE lifetime and was found to be within a few percent of the velocity of the latter two elements. The period considered here extends from July 2007 to November 2009, during a time when the monthly average sunspot number was not larger than 10 for any two months in a row. This time is well matched with the minimum of solar cycle 24. During this period of time, polar coronal holes were always present and rather stable, thus the solar rotation determined the solar wind type observed by ACE, alternating it between fast and slow solar wind (see, Schwenn and Marsch [235])

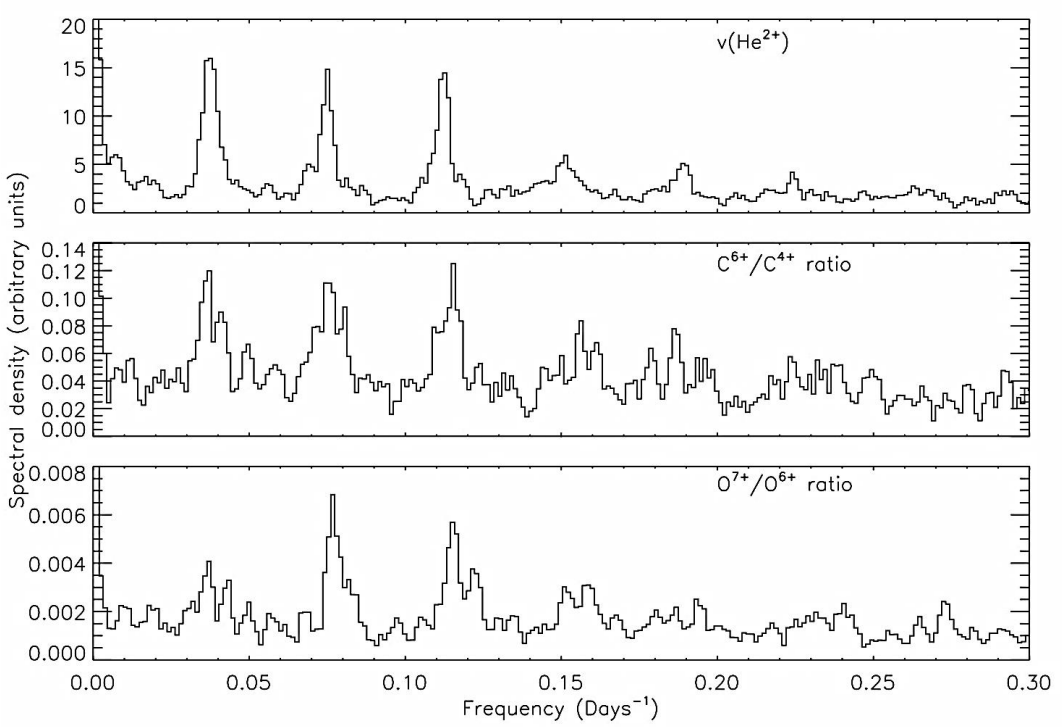


Figure 43. Fourier power spectrum of: the solar wind velocity (measured as the bulk speed of He^{2+}; top), the C^{6+}/C^{4+} ratio (middle), and the O^{7+}/O^{6+} ratio (bottom). Peaks correspond to the synodic solar rotation rate and its overtones. (Figure 43 is a direct reprint of Figure 3 from Landi et al. 2012)

The power spectrum of the He^{2+} speed clearly shows a few well-defined peaks consistent with the frequency of the solar rotation rate and its first five overtones, corresponding to a rotation rate of 27.5 ± 3.0 days. The uncertainty is taken as the Full-Width at Half-Maximum (FWHM) of these peaks. After conversion to a sidereal rotation rate, this rotation can be compared to estimates made by Snodgrass & Ulrich [236] and Newton & Nunn [233]. We found that it corresponds to the rotation rates of photospheric features rooted at latitudes lower than 45°. The middle and bottom panels show the power spectrum of the C^{6+}/C^{4+} and O^{7+}/O^{6+} ratio, respectively, and both show peaks corresponding to the same frequencies as He^{2+} speed, although they are somewhat broader; those corresponding to harmonics higher than 3 are buried in the noise. The C^{6+}/C^{4+} fundamental frequency indicates a slightly lower rotation rate (26.5 ± 4.1 days) than the O^{7+}/O^{6+} frequency (27.3 ± 5.2), but the differences are well within uncertainties; however, the frequencies of the O ratio overtones tend to be systematically larger than those of the C^{6+}/C^{4+} and He^{2+} speed. Also, the peak of the fundamental frequency is larger in the C^{6+}/C^{4+} power spectrum than in the O^{7+}/O^{6+} spectrum. If we interpret this periodicity as the effect of systematic changes in solar wind type and composition over the time-period of many solar rotations, **Figure 43** clearly indicates that the sensitivity of the C^{6+}/C^{4+} and O^{7+}/O^{6+} ratios are comparable, but that C^{6+}/C^{4+} is less noisy and better resolves rotation-dependent signatures.

5.1.6. **Discussion**

This finding marks a significant contribution to the heliospheric science community, as the O^{7+}/O^{6+} ionic charge state ratio was used as the primary indicator of solar wind type and source-region for over a decade [237]. The bulk quantity of O^{7+} is very small compared to that of O^{6+} for most of the temperature range (leading to statistical uncertainty), however, C^{4+} and C^{6+} ions have similar relative abundances [230], and thus are more easily measurable (leading to higher statistical certainty). This is the primary reason that the harmonic signatures were more clearly resolved in the spectrum of C^{6+}/C^{4+} than O^{7+}/O^{6+}. **Figure 44** displays the variance of C^{6+}/C^{4+} and O^{7+}/O^{6+} charge states across a range of solar wind temperatures.

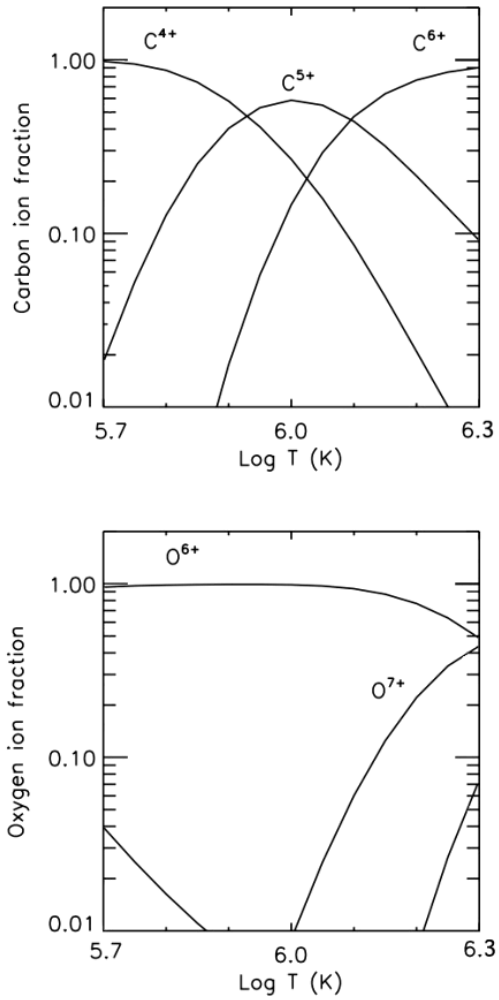


Figure 44. Ionization fraction of carbon (top) and oxygen (bottom) as a function of temperature. The *x*-axis spans the typical range of solar wind temperatures. The C^{6+}/C^{4+} ratio allows the sampling of a larger temperature range than O^{7+}/O^{6+}. (Figure 44 is a direct reprint of Figure 3 from Alexander et al. 2011)

While the evaluation of the C^{6+}/C^{4+} power spectrum could have been conducted through a purely visual scan, the SWICS instrument produces data for 40 individual ionic charge states, and audification provided an efficient method for comparative analysis across spectra. Additionally, the error margin for the synodic solar rotational period as derived by the FWHM of the Fourier analysis (26.5 ± 4.1 days) was in the 10–20% range, while the error margin of the rotational period derived through auditory analysis (26.94 ± 0.20 days) was less than 1%. In this case, accuracy provided through auditory analysis marked an order-of-magnitude improvement over traditional analysis methods. Furthermore, the ability to detect subtle differences in frequency is not necessarily

limited to the skillset of an audification specialist, as research indicates that musicians and non-musicians are able to reliably discriminate between subtle frequency differences of less than 1% [238].

Two additional points can be drawn from this case study: **1)** The audification specialist needed a sufficient level of domain-specific knowledge and familiarity with the data to recognize the harmonic signature of C^{6+}/C^{4+} as a significant observation worthy of presenting to the group, and **2)** the ensuing physical investigation required the expertise of the SHRG, hence the significance of the initial observation could not have been appropriately situated within a broader scientific context without the knowledge of domain scientists. A certain level of immersion was necessary for the audification specialist to gain a basic level of fluency; however, a full understanding of the mathematical methods applied by the SHRG was not expected. In this respect, the delineation of working roles was clear.

It should be noted that this type of purely exploratory research is rare within the heliospheric sciences, where even highly experimental work is traditionally directed by a clear research hypothesis (whether stated or unstated). At the outset of the investigation the domain scientists proposed a parameter mapping approach that would "tell the story of the solar cycle in acoustic terms." In this way the audification specialist was encouraged to create an artistic and inherently musical representation of the solar wind. Toward this end, various suggestions were provided including the creation of a percussive element to track the 27-day solar rotational period. The resulting parameter mapping sonification included renderings of solar wind density, velocity, and charge states, and was well received at the 2010 International Conference on Auditory Display. In the conference proceedings the specialist noted that members of the SHRG were able to auditorily track interactions between multiple data parameters, but "had yet to unearth any new findings from initial experimentations" [239]. While the early parameter mapping experiments did not produce a domain-specific scientific outcome, they successfully established a working relationship that catalyzed ongoing cross-disciplinary collaboration. As the audification specialist began meeting with the SHRG on a weekly basis the primary focus of the investigation shifted to practical data-mining applications for auditory analysis. This work was guided by a single research question: *can*

sonification reveal new features in the SWICS data archive? In pursuing this question, the specialist generated the first audification from SWICS time series on Wednesday January 12th, 2011. Over twenty solar wind data parameters were audified and assessed over the span of several hours, and all audio files were found to contain some level of rotationally dependent "hum." The auditory analysis of C^{4+}, C^{5+}, and C^{6+} ionic charge states was presented in a meeting with the SHRG the following day. The rapid insight gleaned through audification was unanticipated, as various parameter mapping experiments had not borne fruit over the previous year. The potential for audification to expedite the process of feature identification in the evaluation of long time series will be explored more deeply in the three following case studies.

5.2. Simultaneous Wave Observations in THEMIS Data

The remainder of this chapter will focus on interactions with research scientists at NASA Goddard Space Flight Center. The domain-specific discovery detailed in this section stems from a follow-up meeting after an initial workflow assessment with a research scientist in the Heliophysics division. Working alongside the scientist, the specialist was able to apply auditory analysis to quickly identify a spectral feature that was found to contain a novel interaction between several wave-modes; a larger investigation was subsequently presented at the 2014 American Geophysical Union (AGU) meeting in San Francisco [223].[29] A brief mention of audification was included in this presentation; however, the role of audification in the analysis process was not explored in any great detail. This section outlines the interaction between the research scientist and the audification specialist, drawing from a transcribed audio recording to explore the role of audification in the task of feature identification.

5.2.1. **Scientific Rationale**

In a workflow assessment it was found that the scientist's research focused primarily on energy transport in collisionless shocks, including bow-shocks and those produced by coronal mass ejections [240]. Specifically, his research focused on properties of the free (excess) energy sources that produced waves, and how these waves

[29] A manuscript is currently in preparation for submission to the journal of Geophysical Research Letters.

affect the macroscopic structure of a shock. In this investigation he routinely worked with data from the Electric Field Instrument (EFI) on the THEMIS satellite [241-243], conducting polarization analysis to identify wave modes (and determine propagation direction). The scientist indicated that he generally avoided any "black box" software, as he preferred to have full knowledge of what was happening during every step of an analysis process. For this reason, he constructed his own software for accessing and analyzing long time series from the THEMIS satellite (e.g., applying wavelet transforms to visualize spectra and generating summary plots of temperature, distribution, and bulk flow).

The original objective of this workflow assessment was to gain a sense for traditional working practices in the heliospheric sciences with the larger goal of applying this knowledge toward the development of a broadly accessible tool for auditory analysis. As the initial meeting with the scientist drew to a close, he noted that some electrostatic waves, known as "electrostatic ion acoustic" waves, oscillate compressively and behave very similarly to sound waves. The specialist inquired as to whether these phenomena had been audified, and the domain scientist was uncertain. This dialogue piqued the curiosity of both parties and initiated and an exploratory investigation into the use of audification for the identification of spectral features in THEMIS data.

The Time History of Events and Macroscale Interactions during Substorms (THEMIS) spacecraft was launched on February 17, 2007. During the three-year prime mission, the collection of five satellites observed both fields and particles, with the stated mission objective to "elucidate which magnetotail process is responsible for substorm onset at the region where substorm auroras map" [242]. Data gathered by the particle and field detectors on the spacecraft also allow researchers to investigate particle beams, shocks, and wave-particle interactions in the solar wind (particularly those effected by sun-earth interactions) [242]. One of these detectors, the EFI instrument, samples the interplanetary electric field at a rate of 16,384 Hz [241].

5.2.2. **Data Selection and Preparation**

As a part of a larger investigation seeking to quantify the energetic contribution of wave-particle interactions at collisionless shocks, the scientist was in the process of surveying THEMIS EFI data for instances in which the spacecraft passed through Earth's bow shock [244]. He noted that as the instrument measured the intensity of the magnetic field at a particularly high sampling rate during "burst mode," it would be possible to generate an audio file with a sampling rate equal to that of the EFI instrument and listen to spectral features at the rate which they naturally evolve in the solar wind (without the application of a temporal scaling factor).

The file initially selected for audification contained several periods of burst-mode observations concatenated into a single array. These data were gathered on August 23rd, 2010, when the spacecraft was near the earth's magnetopause. In preparation for analysis in a scientific computing platform the data had been passed through a single pole high-pass filter with a cutoff frequency of 10 Hz [245]. This removed any DC-offset and subsequently allowed for the creation of an audio file with amplitude values centered about the zero crossing (i.e. the resulting waveform naturally tended to center between minimum and maximum amplitude values). In this way the data closely emulated a traditional audio recording and thus provided an ideal candidate for audification.

5.2.3. **Audification**

Time series observations gathered by the EFI instrument at a sampling rate of 16,384 samples per second were imported into an IDL array. The *WRITE_WAV* function was used to write a 16-bit audio file with the same sampling rate. The original IDL array contained 1,150,976 samples and the resulting audio file was approximately 1 min 10 sec in length.

An initial auditory scan revealed significant digital distortion during playback. This was attributed to a clipping procedure that had been applied to truncate large amplitude spikes and maximize the signal to noise ratio in the resulting audio file. The minimum and maximum allowable range for this clipping procedure was expanded, which reduced the overall gain of the audio file and diminished the distortion. Any remaining clipped values were assigned as NaN's (along with any bad or missing data)

and linearly interpolated to further eliminate any artifacts introduced by the audification process.

The scientist originally wrote each of the three magnetic field components (x, y, and z) to individual channels in a single audio file, but the format proved unwieldy and unreadable by many common media players. Subsequently, each component was written to an individual audio file such that these files could later be processed and spatialized using third-party software.

5.2.4. **Auditory Analysis**

The audified x-component was imported into iZotope Rx for visual and auditory analysis. After an initial listen through the entire file, the domain scientist stated that the data sounded "weird" and not at all like the generic "noise" he expected; the specialist observed the audification to be highly featured, quite unlike observations of the solar wind gathered at lower sampling rates (which are dominated by omnipresent broadband noise).

The file began rather quietly with no significant structural features aside from the subtle presence of the instrumental noise floor (the specialist remarked on the high signal-to-noise ratio). As the original array contained several concatenated data sets, the resulting audio file displayed a succession of distinct regions, each with a unique "family" of spectral signatures. For instance, the first 12 sec contained a persistent low-frequency (> 50 Hz) rumbling and an occasional high frequency "chattering" that resembled faint bird-song. The following 12 sec were relatively quiet and punctuated by narrow-band noise signatures sweeping between high and low frequencies (approximately 300 Hz – 2 kHz). The following segment contained a persistent "scratching" sound (similar to that of a small animal furiously burrowing in a pile of dry leaves), as well as a faint low-frequency hum. The "scratching" intensified in the second half of the audio file and in several places resembled the sound of small rocks falling on a thin metallic sheet. This segment was approximately 30 dB (RMS) louder than the initial quiet section.

In the investigation of these spectral features the specialist initially focused less on broadband noise activity and more on several sporadic incidences of narrowband

resonance. The second half of the file contained a narrow bandwidth of increased spectral power drifting between ~100–300 Hz that occurred in the presence of impulsive high-frequency noise bursts. Auditorily, this feature resembled a slowly rising and falling resonance that one might describe as wind eerily blowing through the trees. The specialist located an isolated incidence of this phenomenon and presented a power-spectrum analysis conducted in iZotope Rx. The scientist readily categorized this feature as a *lion roar*—a type of electromagnetic fluctuation that propagates close to the background magnetic field, detectable in Earth's magnetosheath at frequencies near a few hundred Hertz. This name was originally applied when the feature was played back over a loudspeaker and the sound resembled the low roar of a lion on the Serengeti [184]. This was the first time the domain scientist had listened to a *lion roar* event, and the etymology of the term was immediately understood.

The specialist shifted focus to a series of resonant narrow-band spectral peaks sweeping between 200–500 Hz. The first event occurred approximately 15 sec into the audio file (corresponding to data sample 243,000). The specialist relayed the precise location to the scientist who proceeded to investigate the frequency spectrum of the EFI data along with a number of associated data parameters. The scientist was initially unable to classify these features, however, several minutes later they were identified as dispersive waves (the rise and fall in frequency was attributed to variation in local electron density).

The specialist proceeded to extract a number of auditory features that were readily correlated with various high-frequency wave modes including whistler and electrostatic solitary waves. After approximately 40 minutes of investigating in this manner the specialist isolated an event that was initially quite subtle before the application of a significant gain increase. This feature contained a transient high frequency resonance that smoothly descended in frequency as it dissipated; the scientist likened its sound to the "chirp" of a bird (see Figure 45).

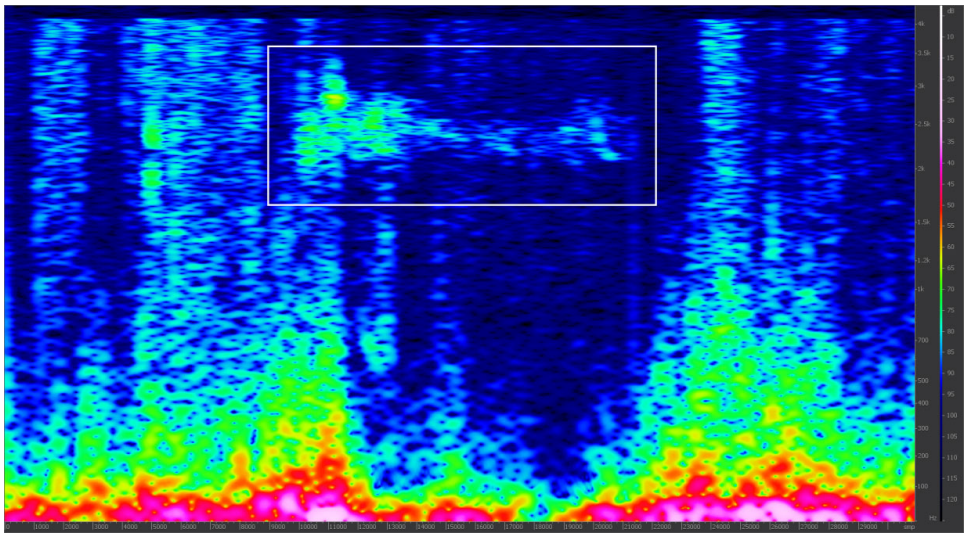


Figure 45. A spectrogram representation of the "chirp" event displayed in iZotope Rx. Here distinct wave modes can be seen as narrow-band high-frequency activity occurring in the presence of strong low-frequency power (see white box).

When asked if he would like to "zoom in" to a specific region of the audio file, the scientist selected a narrow window containing a band of increased spectral power between 2–3.5 kHz. He then noted the event time and proceeded to investigate the corresponding EFI time series in software specialized for scientific computing. At this time the following exchange took place between the audification specialist (AS) and the domain scientist (DS):

AS: There's a kind of double peak that happens.

DS: I'm not sure I believe it though.

AS: It sounds real.

DS: It's not, it's aliased… your data is under sampled… you can't really say much about it.

AS (While viewing the waveform in iZotope Rx): It looks like there should be enough data points to represent that waveform when I zoom in and look at the individual data samples.

DS (While viewing the waveform in scientific analysis software): Look at them… they're spiky, that means it's under-sampled, if it were well resolved it'd be smooth… *pause* Oh no I forgot… all this processed data that I rotated into a different coordinate system, the software I use automatically down-samples it to the

search coil data time stamps because… sometimes you want to compare the search coil to the electric field… so I reduced the sample rate from 16,000 to 8,000… that's why you're seeing it nice and smooth.

In the previous exchange the two were observing the full-resolution EFI data loaded into iZotope Rx on the specialist's laptop, and the down-sampled data loaded in scientific computing software on the scientist's computer. The scientist proceeded to load the raw data (at the full sampling rate), but was unable to generate a corresponding spectral plot without additional time-intensive processing. The following section documents the ensuing investigation of this feature, which was found to contain a novel simultaneous occurrence of several wave-modes.

5.2.5. **Knowledge Extraction**

The "chirp" was located roughly 94,000 samples into the original data file and spanned approximately 11,000 samples. Visual inspection of the time series in iZotope Rx initially revealed high-amplitude low-frequency fluctuations and closer inspection uncovered periodic activity at higher frequencies. The scientist was able to quickly identify bi-polar pulses indicating the presence of electrostatic solitary waves, one of the most important wave-modes in plasma physics [246]. A higher frequency signal was superimposed over this wave-mode, occurring below the local electron cyclotron frequency of 3 kHz. The frequency of these waves was found to be in direct correlation with the intensity of the background magnetic field, such that an overall decrease in field strength correlated with a decrease in frequency. This observation—along with the analysis of several related data parameters—led the scientist to uncover the simultaneous occurrence of lower hybrid and electromagnetic whistler mode waves (lower hybrid waves occur near the local lower hybrid resonance frequency, which is roughly 43 times lower than the electron cyclotron frequency for the events discussed here). The simultaneous coupling of these wave modes had been predicted in theory and simulation but never observed in nature [247-250], and it is this unique coupling that produced the "chirp" sound originally identified in the sonification.

These data were passed to a researcher at the University of Minnesota who conducted an extensive investigation of the plasma and found large fluctuations in

density and temperature. This time period was ultimately found to contain four simultaneous distinct wave modes: lower hybrid, whistler, electrostatic solitary, and electrostatic electron cyclotron harmonic waves. Bi-streaming electron beams were identified, which could potentially provide the free energy for the instability predicted to produce the observed waves. It was concluded that the results of this analysis "provide insights into wave coupling near the magnetopause and suggest that coupling processes may be more important than previously thought" [223]. These findings were presented at the 2014 AGU conference in San Francisco [223].

5.2.6. **Discussion**

The structure of this collaborative investigation was quite unlike the previous exploration of ACE SWICS data, in which the specialist conducted the auditory analysis before approaching the SHRG with a concise set of observations. In this case the specialist had no prior experience working with the highly featured data from the EFI instrument and immediate feedback from the domain scientist enabled the rapid creation of a catalogue of spectral features.

Revisiting the transcript—in the moment that the domain scientist expressed disbelief as to the legitimacy of the "chirp" feature, the audification specialist made an objective judgment based purely on auditory assessment in stating "it sounds real." The specialist had limited exposure to a small representative sample of data from the THEMIS satellite, hence this statement was not so much a comparative evaluation across spectral features found in the EFI data as it was a qualitative assessment based on previous observations of instrumentally induced noise and sampling artifacts encountered in a variety of audified data sets. Subsequently, the domain scientist expressed skepticism that the time series produced by the EFI instrument had sufficient resolution to accurately render the high-frequency "chirp" event, that is to say—the data were *aliased*. "Aliased" is a signal processing term commonly applied in the fields of space science and audio recording to indicate a specific type of distortion that occurs when a continuous signal is digitized with a sampling rate that is insufficient to properly resolve a relevant signal [251]. The specialist, familiar with this terminology, understood that a visual inspection of individual data points on the audio waveform would be the quickest way to determine

whether aliasing was indeed present (as auditory evaluation suggested it was not). It was this inspection that prompted the domain scientist to conduct his own visual analysis of the time series, at which time he remembered a crucial step that had been overlooked in his analysis pipeline.

The "chirp" feature initially eluded detection for several reasons: **1)** The intensity was 30 dB lower than other, more prominent features found later in the same data file, **2)** the feature was relatively short in duration and the high-frequency waves visible in the time series were overshadowed by high-amplitude low-frequency activity, and **3)** the scientist had yet to prepare a spectrogram representation for the full-resolution (16 kHz) data. The results of the studies presented in Chapters III, IV and VII support the claim that audification can successfully reveal subtle spectral features in time series data, this case goes one step further in demonstrating how the tools of audification may be applied to rapidly evaluate the quality of not only the resulting frequency spectrum, but also of the original time series. In assessing for the presence of aliasing, the specialist and scientist gleaned new insight upon closely examining the audio waveform as a two-dimensional line plot. This was possible only due to the isomorphic nature of the audification process, and in this specific case, the first audio waveform assessed by the researchers had a higher temporal resolution than the first line-plot produced by the domain scientist.

This investigation was purely exploratory, as the scientist had not previously expressed any interest in investigating data through audification. However, after he noted that many features within his data naturally fell within an audible range, he was quite surprised to learn that most scientific computing packages had built-in functions for data audification. He then took immediate initiative in exploring the WRITE_WAV function in IDL. While he did not continue to routinely apply audification in the analysis of THEMIS data, he did provide a number of suggestions for novel audification use-cases for the RBSP and Van Allen Probe missions.

5.3. Multimodal Identification of an Extended Proton Cyclotron Wave Storm in *Wind* MFI Data

This section provides a scientific result from the work presented in Chapter IV, in which a research scientist at NASA GSFC was able to apply audification methods to identify a region of *Wind* MFI data that was found to contain wave-particle interactions. In a series of interviews the scientist indicated that there was an overwhelming volume of time series data within their subfield, and many experts relied on time-intensive visual scanning methods for feature identification. In lieu of visual assessment, the scientist would often apply automated search routines, though he found it "impossible (for me) to check if it worked on all the data… because it's just so huge that I can't really look through." Audification was proposed as the high sampling-rate could allow for rapid scanning of an entire year's worth of data.

The two previous case studies began as purely exploratory investigations that bore fruit in novel observations. In contrast, this case study presents work in which multimodal analysis was conducted with the specific intent of identifying regions that potentially contained wave-particle interactions. The audification specialist reviewed *Wind* MFI data from 2008 and presented a list of candidate features to the scientist for additional analysis through traditional methods. One particular region identified as "the mother of all whooshes" was found to contain an ion-cyclotron wave storm (ICWS) associated with an abnormally dense proton beam occurring on the 4th of November 2008. This case study provides an example in which audification was applied to survey a large archive of time series data, resulting in the identification of a new feature of scientific interest.

Sections 5.3.1 to 5.3.5 are adapted from R. T. Wicks, R. L. Alexander, M. Stevens, L. B. Wilson III, P. S. Moya, A. Viñas, L. K. Jian, D. A. Roberts, S. O'Modhrain, J. A. Gilbert, and T. H. Zurbuchen (2015), *A proton cyclotron wave storm generated by unstable proton distribution functions in the solar wind* [207], which is currently under review with the Astrophysical Journal. The section headings have been revised to match the format of this chapter. The audification specialist identified the event discussed in this manuscript, and significantly contributed to the writing.

5.3.1. **Scientific Rationale**

The solar wind is only weakly collisional, therefore wave-particle interactions are important in determining the evolution of the proton distribution function. Energy can transfer from fields to particles and vice-versa. Field-to-particle transfer of energy is important in explaining the non-adiabatic heating of the solar wind [252], the anisotropic shape of thermal particle distribution functions [253], and the dissipation of large scale turbulence [254]. Particle-to-field energy transfer may arise from unstable particle distribution functions [255-260], for example as a product of large scale turbulence [261], shock acceleration, or reconnection, and this transfer generates kinetic plasma waves [260, 262]. These fundamental processes are determined by small-scale interactions, thus we need to observe magnetic and electric-fields at ion-kinetic scales and frequencies, such as the proton gyro-radius ρ_p and gyro-frequency Ω_p, in conjunction with detailed observations of particle distributions.

There are numerous different types of plasma waves that may interact with the proton distribution to exchange energy. The most commonly observed coherent waves close to ion scales in the solar wind are ion-cyclotron waves, which can be found in individual wave packets lasting just a few minutes [217], or in "cyclotron wave storms" lasting many hours [263]. Cyclotron waves seen in the solar wind frame are left-hand polarized electromagnetic plasma waves with frequencies close to the proton gyro-frequency and wave vectors often close to the local magnetic field direction (quasi-parallel). Surveys of STEREO and MESSENGER spacecraft data have been used to identify and study ion-cyclotron-wave storms [263, 264], although a complete description of how and why such storms happen is currently lacking.

Here we use the *Wind* spacecraft to study kinetic plasma waves. *Wind* provides an unrivaled data set containing nearly 20 years of solar wind observations with magnetic field measured at a 0.092 sec sampling rate by MFI [201] and thermal particle distributions measured every 92 sec by the SWE instrument [265]. In order to identify regions that exhibit wave-particle interactions we look for magnetic-field fluctuations that display properties similar to those expected for proton-kinetic plasma waves. The challenge is to do this in an efficient way since observations gathered at a 0.092 sec sampling rate provide 950,400 samples per day and more than 6.59×10^9 observations

over the lifetime of the *Wind* mission so far. The waves we are searching for typically have a period of a few seconds so scanning through the data visually with sufficient resolution to observe the waves would require many person-months of effort.

5.3.2. **Data Selection and Preparation**

The *Wind* spacecraft magnetometer (MFI) [201] provides 11 Hz vector magnetic field observations nearly continuously for the entire mission lifetime. We chose to study the year 2008 because it is during the recent solar minimum which should limit the number of complex events occurring in the year. During this period the *Wind* spacecraft was orbiting near the L1 Sun-Earth Lagrange point and was continuously immersed within the solar wind plasma flow.

5.3.3. **Audification**

Audification allows us to survey the entire magnetic field data set for the year 2008 in a relatively short time and compare the events found easily. At a typical audio sampling rate of 44,100 samples per second, 66.8 minutes of magnetometer observations at 11 Hz lasts just 1 sec in audio playback. The typical human ear can perceive sound between frequencies of 20 Hz and 20,000 Hz with maximum sensitivity between 2,000 Hz and 5,000 Hz. This means that for an 11 Hz time series sampled into audio at 44,100 samples per second, frequencies between 0.005 Hz and 5.5 Hz are audible, with maximum sensitivity between 0.5 Hz and 1.25 Hz. The range of audible frequencies therefore typically includes the inertial range of turbulence, Ω_p, and ρ_p. The region of maximum auditory sensitivity is close to the range of frequencies at which we expect the cyclotron waves to occur, making the method particularly useful for identifying proton-kinetic waves.

Daily data sets of 11 Hz magnetometer observations from the *Wind* spacecraft gathered over 2008 were acquired from the SPDF Coordinated Data Analysis Web archive and converted into 12 audio files spanning one month of data each. We chose the $\hat{z}$ GSE component of the vector for auditory analysis as it had the lowest amount of spacecraft spin induced noise. One year of data for this single vector component contains

roughly 330 million data points, and the resulting audio file is approximately two hours in length when played back at a rate of 44,100 samples per second.

5.3.4. **Auditory Analysis**

The resulting audio files were then assessed both auditorily and visually in the iZotope Rx software platform. iZotope RX generates a time series and spectrogram representation of the data which is scrolled through as audio playback occurs, a regular short-time Fourier transform with a window size of 2048 samples is used in the creation of the spectrogram. The turbulent solar wind presents itself auditorily as highly compressed broadband noise with modulations in amplitude occurring at both small and large time scales, sounding somewhat like a flag waving in a strong wind. An example of an audio file containing the audification of the B_Z GSE component of the *Wind* MFI data from a relatively unremarkable region of turbulence is available for listening online [208].

Through auditory and visual analysis using iZotope Rx 410 regions were identified as potentially containing features of interest. A subset of 7 regions (with an average length of 6 seconds in the audio file) was selected for additional analysis. The 4th of November 2008 (DOY 309) was selected as the best candidate for this case study due to activity that manifested as a long-duration, intense whooshing noise in the audified data, the audio file from this day is available for listening online [208].

5.3.5. **Knowledge Extraction**

Figure 46 shows a summary of the solar wind conditions during the November 4th 2008 (DOY 309). For most of the day the solar wind magnetic field is very radial and points back towards the Sun. During this interval a dense proton beam occurs, drifting close to the Alfvén speed ahead of the core of the distribution. Simultaneously waves are seen and heard in the magnetic field data, the waves have amplitudes close to 0.3 nT and, although intermittent, last for more than 8 hours.

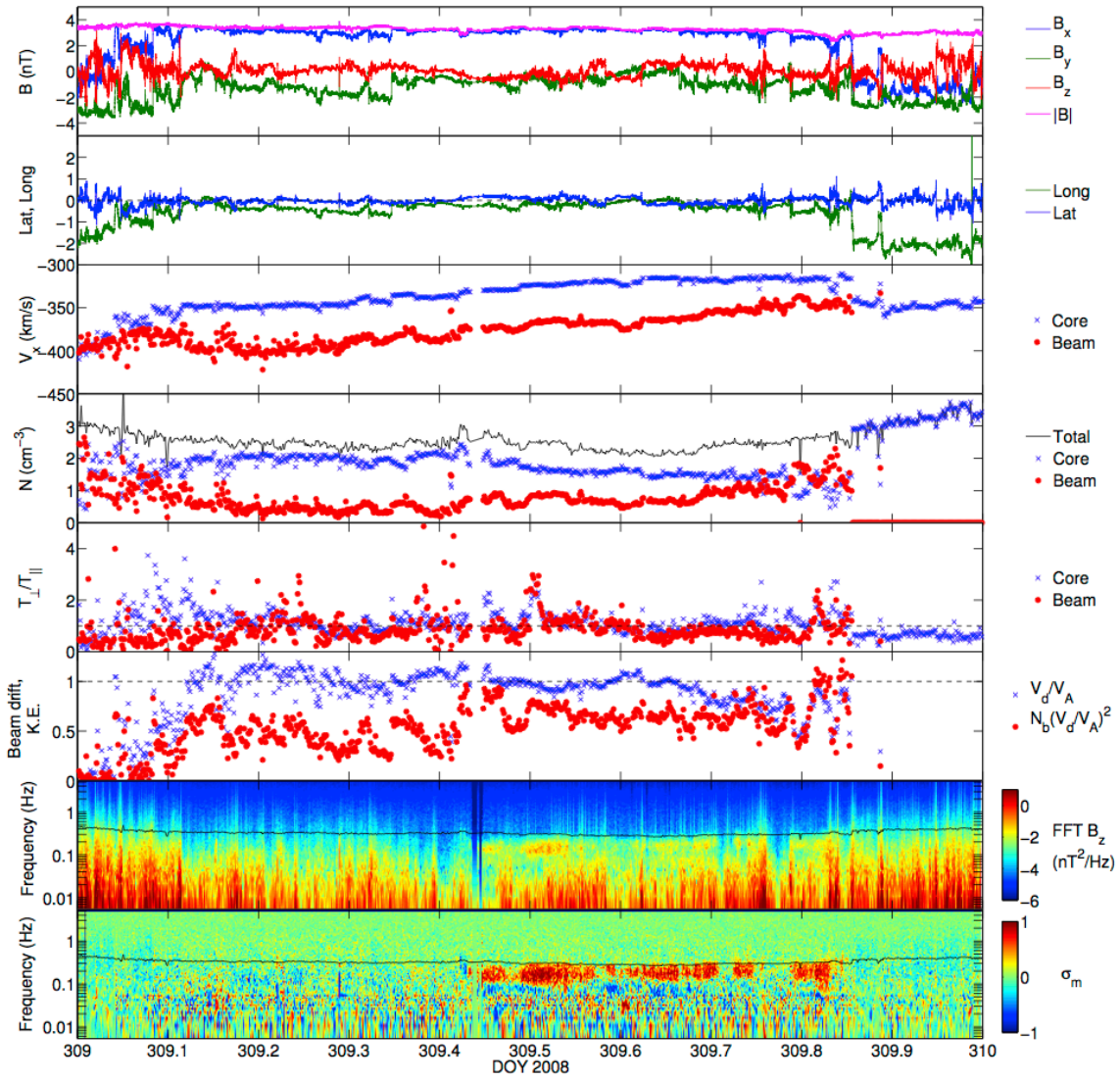


Figure 46. Time series plots of the observed magnetic field components (top panel) and angles (second panel), the proton core and beam velocity (third panel), density (fourth panel) and temperature anisotropy (fifth panel), proton beam drift speed relative to the core normalized to the Alfvén speed and the related kinetic energy (sixth panel) and the spectral properties of the magnetic field, the Fourier power spectrum of the B_Z component (seventh panel) and the reduced magnetic helicity (bottom panel). The wave activity can be seen clearly in the bottom two panels and correlates with high temperature anisotropies ($T_\perp/T_{||} > 1$) and dense and energetic beam distributions. Figure 46 is a direct reprint of Figure 1 from Wicks et al. (2015).

5.3.6. **Discussion**

In this case study, auditory analysis was conducted within a software platform that provided simultaneous spectrogram visualization, enabling rapid audio-visual exploration of a large time series that would be considered quite daunting within most platforms dedicated to scientific computing (e.g., Matlab and IDL). This event, originally identified by the audification specialist as "the Mother of All Whooshes… an extremely long and very clear whooshing," was immediately recognized by the domain scientist as a potential feature of interest due to the extended presence of high-amplitude wave activity. The ensuing literature review revealed that this precise region had been previously examined, though the investigation primarily focused on the release of solar energetic particles and did not survey the associated magnetic field data [266]. While it is true that many events uncovered through multimodal analysis could also be identified through visual scan, here we find an example of a feature that had not been observed primarily due to the vast archive of available data. Not only does the MFI instrument produce hundreds of millions of magnetic field observations per year—there are a vast number of instruments across numerous sun-observing satellites that produce a wide range of data parameters on an hourly basis (e.g., proton velocity and density, ionic charge states).

The initial analysis conducted by the audification specialist unearthed a number of features with exceptionally complex spectra, often bordering the Nyquist frequency of the MFI instrument. In some instances the domain scientist was able to quickly attribute these features to time regions in which the *Wind* satellite passed through Earth's magnetopause (a region rich with high-amplitude wave-particle interactions, well documented in the literature). In other instances he noted that these anomalous features would require extensive investigation in order to understand the underlying physical phenomena. Focusing the initial analysis on a single long and clear event was a measured decision on the part of the scientist, who indicated that "if this new analysis technique shows me the things I know and it shows me something new then I'm more likely to believe the 'something new.'" While he had not previously observed the event on the 4th of November, he was quickly able to determine that the "whoosh" was generated by ion cyclotron activity. The ensuing investigation into the source of this activity grew quite complex, requiring theoretical and observational analysis from multiple collaborators at

NASA GSFC and MIT; however, this distinct event was clearly resolved in the MFI data and presented a well-bound physics problem.

Closer examination of the evolving work-practice points to a crucial element of the collaboration between the domain scientist and the audification specialist: features that were immediately recognized as auditorily interesting were often not entirely novel. Extremely dramatic spectral signatures were likely to be well documented, while subtle, more nuanced features revealed through audification were often difficult to investigate through traditional analysis methods. The audification specialist learned to readily draw upon the expertise of the domain scientist in order to quickly reject auditory signatures that were subtle and often erroneous, thus expediting the analysis process. For this, it was helpful for the specialist and the scientist to be physically located in close proximity, such that interactions could happen fluidly and spontaneously. On numerous occasions, a discussion was prompted when a dataset was played back over speakers, and this held true even when the researchers were not initially working in close physical proximity.

The domain scientist also displayed strong interest in adopting multimodal analysis methods. He not only downloaded the iZotope Rx software and conducted independent analysis, he also wrote custom parameter mapping algorithms in Matlab to sonically overlay relevant data on the audification. More specifically, he calculated an important value known as the "gyro radius," and mapped this parameter to the frequency of a sine wave that was then superimposed on the audified solar wind data. The scientist noted that this experiment was unsuccessful due to the effects of auditory masking; however, the audification specialist was able to glean new insight as to the high variability of the data through listening to the reference tone.

An exit interview was conducted upon the conclusion of the case study. The domain scientist noted that the audification preserved a "richness" in the data that was lost through visual analysis. The audification revealed an abundance of small-scale spectral features that he had not anticipated based on previous research, and he went as far as to say, "I'm going to change my whole method based on that knowledge." This suggests that while audification may primarily play a role in feature identification, insight gleaned through auditory observation can inform decision-making at every step of the data analysis process.

5.4. Multimodal Assessment of LFW Storm Activity in *Wind* MFI Data

The previous case study confirmed that multimodal analysis can be successfully applied to identify Low-Frequency Wave Storm (LFWS) activity within high-resolution *Wind* magnetometer data. In this final case study, audification was applied toward the creation of a comprehensive catalogue of LFWS events occurring in *Wind* MFI data during the year 2005, a record spanning 0.34 billion data points. The domain scientist presented the outcome of this investigation to the 2014 Committee On Space Research in Moscow (COSPAR 2014) [224] and an abstract has been accepted to the 2015 European Geosciences Union (a manuscript submission to the Journal of Geophysical Research is currently in preparation). The scientist indicated that she had a higher degree of confidence in the multimodal assessment than the automated search routine that had been applied to identify LFWS events in a similar data set, as the successful identification rate was roughly 4 to 5 times higher through multimodal analysis.

5.4.1. **Scientific Rationale**

Cataloging LFWS activity is a task particularly well suited for audification. Jian et al. (2009) noted that many physicists assess the turbulent solar wind spectrum through plots that use power spectra calculated over hours (and potentially longer time scales), while a survey of Ion Cyclotron Wave events (ICW – A specific type of low-frequency wave) in STEREO data found that they generally occur with a median duration closer to 50 sec [217]. The difficulty in classifying these events lies in their diverse nature (ICW events can vary widely in frequency, intensity, duration, and spectral bandwidth) and the inherent complexity of large time series generated through measurements of the solar magnetic field (which is largely turbulent and variable in nature). For these reasons, LFWS events can be particularly difficult to identify through automated search routines.

To date, assessments of ICW activity in STEREO and MESSENGER data have been produced primarily through visual inspection [217, 264], and a full catalogue of LFWS activity in STEREO data is available for the year 2008 [263]. This case study provides the first catalogue of LFWS events in *Wind* data, which is important for statistical comparison, and is novel, as *Wind* is situated at L1 and provides excellent information for near-earth conditions. The scientist noted that *Wind* also provides

"relatively high-cadence plasma data, including ion temperature anisotropy and differential flow between core and beam proton populations, which are useful for examining the generation mechanism of local instability."

5.4.2. **Data Selection and Preparation**

As with previous case studies, the *Wind* MFI instrument is selected for the high sampling rate, vast data archive, and variety of spectral features. While previous investigations of LFWS activity had focused on 2008, we selected the year 2005 as it was in the declining phase of solar cycle 23, hence there was a lower proportion of CMEs and other short-lived events. The z-component was selected for auditory analysis due to the low prevalence of instrumentally induced noise.

5.4.3. **Audification**

Daily data sets were downloaded from CDAWeb [267] and a custom algorithm in Matlab was used to write 12 audio files (16 bit, 44.1 kHz sampling rate, *.wav* format), each spanning one month of data. Bad and/or missing data values were smoothed via nearest-neighbor interpolation, and the amplitude for each audio file was normalized based on the highest and lowest data values for the given month. The entire year (344 million data points) spanned approximately 2.2 hours of audio playback at a rate of 44,100 samples per second, and each audified data file averaged 10.8 minutes in length.

5.4.4. **Auditory Analysis**

The auditory analysis was conducted by an audification specialist who had accumulated experience working with audified *Wind* MFI data over the course of several months. The analysis was conducted with ATH-M50 professional studio monitor headphones. The audio files were imported into iZotope Rx 3 for playback with tandem spectrogram visualization and all features were labeled as regions and exported in *.txt* format.

Working closely with research scientists, the audification specialist learned to identify a wide variety of common features including those listed in **Table 8**. An audio example for each feature is provided online [208].

Table 8. Features commonly observed in audified *Wind* MFI data played at 44.1 kHz.

Name	Spectral Signature	Time Series Signature	Auditory Signature
Low Frequency Wave Storms (LFWS)	Regions of enhanced spectral power occurring between 300 Hz and 1500 Hz in the audification, often co-occurring with a decrease in broadband spectral power.	Sinusoidal oscillations.	A "whooshing" sound lasting anywhere from a fraction of a second to several seconds.
Mirror Modes	Regions of broadband noise, occasionally occurring before a shock event.	"One sided" spikes in amplitude.	A sudden distortion with a distinct "flavor." This distortion may sound somewhat artificial.
Wave Packets	Transient regions of enhanced narrowband activity, often occurring in close proximity to a shock.	Sinusoidal oscillations.	A "Whistling" sound, often sharply decreasing in frequency over time.
Shocks	Regions marked by a sharp rise and fall in broadband spectral intensity.	A potentially instantaneous jump in amplitude.	A transient percussive event that may sound like a sudden "thud," gun shot, or "knocking."
Warbles	Similar to LFWS events, with more variability in the background field, often occurring in the presence of broadband turbulence.	Intermittent sinusoidal oscillations with high variability.	Similar to a Low Frequency Wave Storm ("whooshing") event but more "crunchy" in nature.

The auditory investigation also revealed several features that had yet to be classified within the heliospheric literature, these features were onomatopoeically defined with names such as "sizzles." Many appeared as wave-particle interactions that only subtly rose above the level of the background turbulence. The most common of these features are listed in **Table 9**.

Table 9. Novel features detected in audified *Wind* MFI data played at 44.1 kHz.

Name	Spectral Signature	Time Series Signature	Auditory Signature
Sizzles	Bands of increased spectral power occurring at high frequencies, often clipped at the upper edge of the frequency spectrum.	High frequency sinusoidal oscillations.	A high frequency "sizzling," sound that may resemble "squishing," or "crackling."
Whips	An extremely short burst of spectral power, usually sweeping upward or downward and found in pairs.	Tight packets of sinusoidal oscillations.	Short "whip" like sounds, occurring in quick succession.

The auditory analysis of the 2005 data took roughly 15 hours to complete. A total of 768 features were identified across the entire year (averaging 64 per month). The average feature length was approximately 62,000 data samples, translating to a duration of 1.4 seconds in the audio file (approximately 1.5 hours in the time scale of the original MFI data). Each feature was provided with a significance index ranging from 1 to 5, where a value of 1 indicated a feature that was only slightly perceptible above the background turbulence and 5 indicated a strong spectral signature prominently distinguishable from background noise.

5.4.5. **Knowledge Extraction**

The audification specialist presented several examples of strong "warble" (WRB) events to the domain scientist, who confirmed they contained regions that fit the criteria for LFWS events. The specialist then assessed the data and concluded that events with a significance index of 3 or lower would most likely be too subtle to meet the criteria for LFWS activity. All LFWS and WRB events rated with a significance index of 4 or 5 (72 events) through multimodal analysis were then compiled into a list and passed to the domain scientist to be assessed for the presence of LFWS activity. The expert was able to assess all 72 events in a single 8-hour session; **Table 10** summarizes the results of this investigation.

Table 10. The number of features identified through auditory analysis versus the number found to contain LFWS activity.

Feature	# Auditorily Identified	# Containing LFWS Events	Percentage
LFWS-5	11	11	100%
LFWS-4	29	23	79%
WRB-5	8	7	88%
WRB-4	23	9	39%
Total	71	50	70%

Applying the criteria from Jian et al. 2014 [263] it was found that while many regions included in the original list did not contain any LFW storm activity, several regions contained multiple events; hence, examination of the 72 regions ultimately

resulted in the positive identification of 159 LFWS events. These events accounted for 0.92% of the total time [224], in excellent agreement with the results of study of STEREO data from 2008 [263].

5.4.6. **Discussion**

As displayed in **Table 10**, the audification specialist demonstrated a 100% success rate in identifying LFWS activity in the 11 regions labeled as LFWS-5, this dropped to 79% in the case of LFWS-4 events. Many of the regions labeled LFWS-4 contained strong spectral peaks and isolated incidences of LFW activity that were not sufficient to meet the criteria of LFW storm. Surprisingly, a larger percentage of WRB-5 events were positively identified as containing LFWS events (88%) than regions labeled as LFWS-4, despite the fact that WRB events had not been associated with Low Frequency Wave storms at the auditory survey was conducted. In this way the broad inclusion of various types of wave-particle interaction in the initial scan was beneficial, as two classes of features that initially seemed auditorily distinct were found to fit the criteria for LFWS events. Only 39% of WRB-4 events, however, were found to contain any LFWS activity. Overall, the successful identification rate across all regions labeled as LFWS-4/5 and WRB-4/5 was 70%, a rate that was considered excellent by the domain scientist, as the individual regions were generally quite short and quick to assess for the presence of LFWS activity.

An automated analysis of magnetic field observations gathered by the MESSENGER satellite in 2008 indicated a 0.2–0.3% occurrence rate for LFW storm phenomena (research conducted by the scientist and colleagues, pending submission to the Journal of Geophysical Research), this is 4 to 5 times lower than the 0.9% occurrence rate traditionally determined through visual analysis coupled with a case-by-case calculation of transverse power, ellipticity, and polarization [263]. However, the 0.92% occurrence rate (in *Wind* data) determined through multimodal scanning in conjunction with traditional analysis methods was only 0.02 percentage points higher. For this reason, the domain scientist felt that validating the multimodal analysis with an automated statistical analysis was unnecessary. Additionally, when asked to compare the results of the multimodal scan with results previously returned by automated scans, the scientist

indicated that she places a much higher degree of trust in the multimodal analysis. When asked to speculate on the low identification rate of the automated search, she stated that "it's risky" to automate such a large number of analysis parameters and noted a number of parameters that are regularly adjusted on the fly (based on expert assessment).

This investigation provides another example in which multimodal scanning played a large role in the initial stage of feature identification. With a large-scale assessment of this nature, feature identification can potentially be more time-consuming than any other step in the analysis process. In this case, the expert indicated that while it took 8 hours to complete the analysis of the 72 regions identified by the audification specialist, it would have taken at least 80 hours to analyze the entire data set from 2005 through traditional analysis methods. It should be noted that the audification specialist had a high degree of familiarity with the data as they had previously spent several weeks assessing *Wind* MFI data from 2008. As the specialist was cataloging all features (rather than focusing solely on LFWS activity), it's possible that a future assessment for LFWS activity could be completed more rapidly.

Moving forward, the scientist has expressed strong interest in applying multimodal assessment to catalogue the occurrence of LFW storm events in other years. This speaks toward the practical utility of multimodal analysis as a tool for feature identification. Of the novel features identified during the multimodal scan of 2005, "warbles" were found to most commonly be associated with LFW storm activity occurring in regions of high turbulence (hence these storms exhibit shorter wave-packets, higher variability, and frequent discontinuities), while "sizzles" have been noted by the scientist as a novel spectral feature worthy of additional investigation ("whips" occur relatively infrequently and hence are of lower interest at this time).

5.5. General Discussion

The four case studies presented in this chapter provide unique insights as to how audification can be applied in the domain sciences for the identification of novel features in large time series data sets. They mark a significant contribution to the heliospheric sciences, and are responsible for several ongoing research initiatives. It should be noted that the primary focus of this discussion is to thoroughly explore common themes that

arose over the course of the four domain-specific research initiatives (Chapters VI and VII will evaluate the effectiveness of a generalized audification method). This section will discuss findings from regular interactions between an audification specialist and domain scientists, revisit each data set from a broad heliocentric context, and investigate the nature of features that were found to be of interest through multimodal analysis.

5.5.1. **Engaging domain scientists**

The audification specialist spent a great deal of time learning domain-specific terminology, traditional working practices (e.g., preferred software and data representation methods), and fundamental scientific concepts. An understanding of the narrative told through several interacting data parameters was important in the process of identifying candidate data sets for exploration through audification. Several years of close interactions with domain scientists also provided the following insights: **1)** In the interest of expediting the auditory translation process, an audification specialist should not hesitate to ask for technical assistance if any difficulty is experienced in acquiring/accessing data and translating files into a suitable format for audification. **2)** An ideal working environment is one in which the specialist understands the best way to go about independently gathering domain specific knowledge (e.g., through textbooks, courses, lectures, instructional videos) and also feels completely comfortable directly approaching research scientists with questions. **3)** Once a baseline understanding of the data is established, the specialist should be prepared to conduct an independent spectral analysis and approach the scientist with a clear set of observations (a spectrogram display or power-spectrum analysis may help facilitate discussion). **4)** Across the four case studies the exploratory investigation was most likely to advance into new territory when the specialist approached a scientist with a specific inquiry driven by multimodal investigation. **5)** As the time scale of the audification will often be far removed from that of the original data, the specialist should translate any auditory observations back into the original time scale of the data.

5.5.2. **Contextualizing the Data**

The exploratory nature of this research is apparent in the transition from one case study to the next, as focus readily shifted from particle observations to measurements of

electric and finally magnetic fields. As the auditory investigation progressed over the course of several years it naturally trended toward the utilization of larger data volumes gathered at higher sampling rates. Though the ACE/SWICS and *Wind*/MFI instruments had vastly different sampling rates, the primary investigation of both included the analysis of approximately one million data samples. In contrast, the investigation of *Wind*/MFI data from 2005 spanned roughly 344 million data points. While the THEMIS/EFI instrument had a much higher sampling rate than *Wind*/MFI, the latter had produced a much larger data archive at the constant sampling rate of 11 Hz.

The case studies in this chapter were presented chronologically, but it is also helpful to view each with regard to the nature of the data that was explored and the window of the broad solar spectrum that was observed. Collectively, these case studies included the assessment of solar data spread across three time scales spanning a range of approximately 10 orders of magnitude. **Case study 1** utilized ACE/SWICS data gathered at a 2-hour sampling cadence; at this rate an entire year of data is compressed into a single second of audio. As discussed, this level of temporal compression translates the 26.94-day solar (synodic) rotational period to an auditory frequency of 137.5 Hz. The full audible frequency range within this audio file spans 20 Hz – 20 kHz (three orders of magnitude), and within the reference frame of the original data this ranges from 4 hours (half the sampling rate of the instrument) to approximately 167 days (i.e., 10^{-7} Hz). **Figure 47** situates this range within the context of the broad spectrum of the solar wind while displaying the transfer of energy cascading from large to small-scale structures. As the amplitude falls off very slowly at larger length scales (lower frequencies), the energy is more evenly distributed across the frequency spectrum of the audified ACE data, and the spectral power distribution of the resulting audio more closely resembles white noise (when played back at the standard sampling rate of 44.1 kHz).

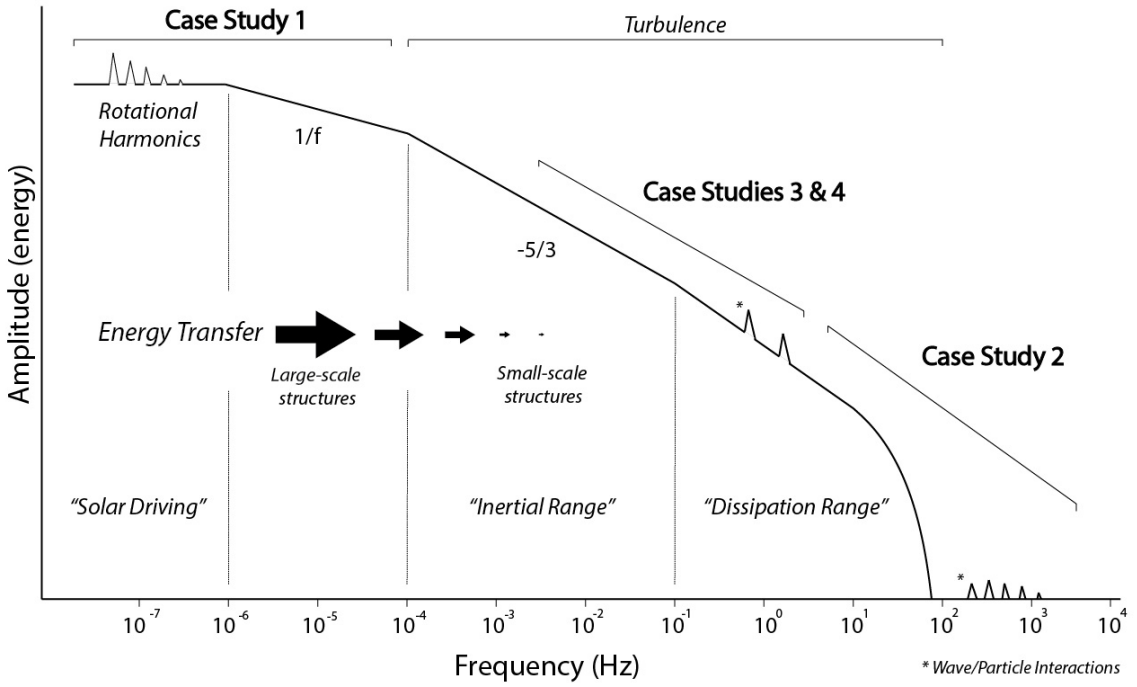


Figure 47. A plot of the energy transfer in the solar wind from large to small-scale structures (low to high frequencies). Generally speaking, energy produced by the sun cascades through "inertial" and "dissipation" ranges before being dispersed through wave/particle interactions. The frequency range encompassed by the four case studies is displayed on the plot. Adapted from Goldstein et al. 1995 [254].

Case study 2 featured data from the EFI instrument on the THEMIS satellite gathered at a sampling rate of 16,384 Hz. As there was no temporal mapping in this instance the frequency spectrum of the audified data fell between 20 Hz – 16 kHz (or approximately $10^1 – 10^4$ Hz in the provided figure), and a full year's worth of data would take one year to listen through (though this data volume is unavailable). The sampling rate of the EFI instrument is approximately 8 orders of magnitude higher than that of the ACE/SWICS instrument, hence the EFI instrument is able to capture micro-scale wave-particle interactions in the solar electric field (the SWICS instrument is ideal for investigating macro-scale structures in solar wind particle distributions). These interactions are largely driven by kinetic energy, occurring within a frequency regime known as the "dissipation range." As the spectral slope is quite steep in this range and wave/particle interactions are intermittent, the high dynamic range of the audified data is understood.

Case studies 3 and 4 utilize data from the MFI instrument on the *Wind* satellite, with a sampling rate of approximately 11 Hz (falling between the respective sampling rates of the SWICS and EFI instruments on the provided figure). At this time-cadence, one year of data requires approximately 2 hours and 18 minutes to listen through completely. This instrument spans the lower bound of the "dissipation" range, and the higher end of the "inertial" range where the properties of the solar wind more closely resemble those of a fluid. The solar wind spectrum drops off with a slope of -5/3, typical of all turbulent spectra occurring in nature. This slope results in a spectrum that closely resembles pink noise, which is slightly easier on the ears than the flat spectrum of the SWICS data.

5.5.3. **Identifying Features of Interest**

While the data in these four case studies encapsulate a broad range of solar activity, the features extracted in each case have notable similarities. All could be described as subtle periodic or resonant behavior occurring in the presence of broad turbulence. While the feature in the second case study occurred in relative isolation, its amplitude fell so far beneath the surrounding turbulent activity that it was essentially "buried" through the normalization process (which would be the case for traditional analysis methods and audification).

In many cases the audification specialist was initially drawn to investigate the most prominent spectral features in the audified data, and these were routinely found to be well documented; in comparison—spectral features that drew the greatest scientific interest were often quite subtle and unremarkable. For example, in the first case study the fluctuations of broadband solar turbulence were extremely pronounced, however, it was the investigation of the relatively subtle evolving harmonic that yielded new scientific insight. Moreover, a handful of pronounced spectral features were assessed in the second case study, all of which were readily classified by the domain scientist; the novel "chirp" event was only uncovered after these features had been catalogued and excluded from the investigation. In the investigation of *Wind*/MFI data it was found that the most striking features were generally caused by explosive events such as coronal mass ejections, while

the energetic fluctuations of LFWS events were several orders of magnitude lower in intensity.

The scientist in the second case study noted that, with reference to THEMIS/EFI data, not many researchers were "willing to get down into the ditch and dig into the stuff to see what's going on." It is indeed a common practice for many heliophysics researchers to generate plots that average power spectra over relatively long time durations. Such representations of the data are excellent for providing a general spectral contour, but any fine-scale structures will be washed out through such a process. In the instance of the third case study, the time period in which the ICWS event was found had already been documented in a separate study, but since the feature under investigation had a much higher amplitude, the waves were never identified.

New instruments currently under development for future sun-observing spacecraft will continue to produce larger data volumes at higher sampling rates. As data archives rapidly expand, a likely shift will occur away from the visual assessment of spectra generated through Fourier and wavelet analysis toward the application of automated search routines that exploit increasingly accessible computational power. Through this process, subtle features in the data may be lost simply because they do not meet pre-established (confirmatory) search criteria. These case studies demonstrate how subtle nuances revealed through exploratory and confirmatory multimodal analysis can produce novel and valuable scientific insights—even in data sets that have been closely examined through traditional analysis techniques. The prevalence of confirmatory analysis practices in the heliospheric sciences has likely consigned a number of important but subtle spectral features to the realm of *unknown unknowns*, a territory ripe for examination through exploratory analysis techniques such as audification.

The following chapter will extend the insights gleaned across the four case studies presented here—establishing a comprehensive framework for conducting auditory analysis through audification, beginning with a flow chart for generating and exploring audified data sets. It will provide researchers with fundamental knowledge to ensure that any audio translation is a faithful and accurate representation of the original data set, and establish an understanding for how decisions made in the audification process effect the audio quality and the temporal progression of the resulting audio file.

CHAPTER VI

Audification Methods

A clearly defined set of methods and best practices are necessary in order for audification to be applied across a wide range of scientific domains as a diagnostic tool for time series analysis. With other forms of sonification, such as parameter mapping, any number of approaches may be appropriated for translating data variables as sound-synthesis parameters, and a great deal of time may be spent investigating novel mapping strategies. There are comparatively few decisions to be made while producing an audio file through audification, and the list of parameters that will have a significant effect on the resulting audio signal can be limited to sampling rate, bit depth, and the number of audio channels used. So long as *reasonable* values for these parameters are chosen, audifications[30] generated by two researchers working with the same data set should sound very similar (and may be perceptually equivalent when matching sampling rates are selected). The range of "reasonable" values will vary depending on the sampling rate of the instrument that gathered the data and the research question posed by the scientist.

This chapter will begin with a focus on these technical aspects of audification, and will subsequently explore other considerations such as how to select an appropriate data set, scale the data, and handle bad or missing data values. Once a workflow has been established, the process of translating data to audio files should be relatively quick compared to the length of time spent manipulating and exploring the audified data. The remainder of this chapter will highlight this process and present several digital signal processing (DSP) techniques that may be used to manipulate the audio in real-time. The material presented in this chapter is modified from the appendix found in R. L.

[30] The term "audification" may be used as a verb when describing the process of directly translating data samples to audio samples, and as a noun when referring to the resulting audio file.

Alexander, S. O'Modhrain, D. A. Roberts, J. A. Gilbert, and T. H. Zurbuchen, "The Bird's Ear View of Space Physics: Audification as a Tool for the Spectral Analysis of Time Series Data," *Journal of Geophysical Research: Space Physics,* 2014JA020025, 2014, pp. 11-12 [205].

6.1. Audification: Step-by-Step Overview

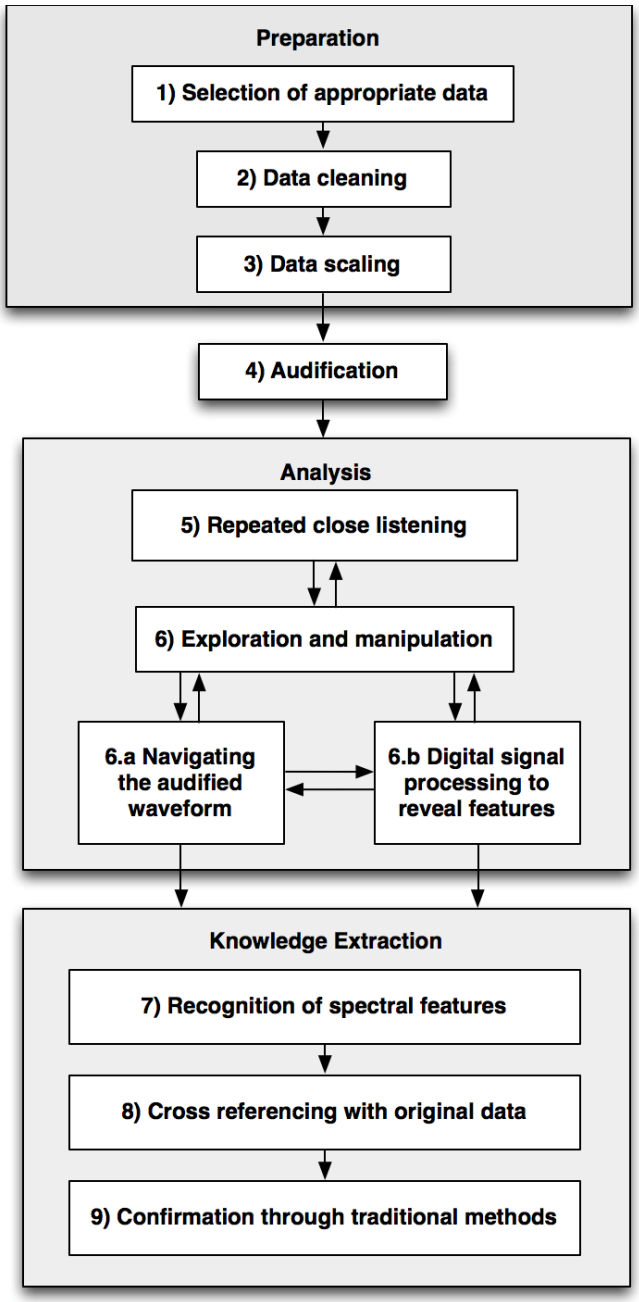


Figure 48. A flowchart of the audification process (adapted from Alexander et al., 2014).

A flow diagram outlining the audification process is provided in **Figure 48**. This standard methodology was iteratively developed over the course of several collaborative research projects with members of the heliospheric science community. These cross-disciplinary investigations were presented as a series of case studies in Chapter V. This methodology is not specific to any software platform, and may be applied in the exploratory and confirmatory analysis of time series gathered from a variety of sources. This chapter provides instructions for data preparation, auditory analysis, and knowledge extraction—expanding upon the framework outlined in **Figure 48**.[31] This set of suggested best practices may serve as a valuable reference for researchers who are interested in applying this analysis technique in the field.

6.1.1. **Selection of Appropriate Data**

The ear can be considered as an extremely sensitive diagnostic tool for the analysis of complex spectra (see Chapter II), and the auditory system is well suited for the detection of subtle and gross features that may appear in the frequency spectrum of time series data sets. Ideal candidates for audification match the following criteria:

1. One-dimensional time series.
2. Observations recorded at equal temporal intervals.
3. A large source of available data (44,100 data samples = one second of audio at the standard rate of sound-file playback).

Data sets may be selected with a clear research hypothesis, toward the goal of extracting a specific class of spectral features (confirmatory analysis); or data may selected for purposes of open-ended evaluation and survey (exploratory analysis).

6.1.2. **Data Cleaning**

Bad and/or missing data entries may produce audible artifacts in the audification. Additionally, values that fall far from the mean can greatly reduce the overall dynamic range of the audio file (resulting in a low level audio signal). **Table 11** offers three methods for processing these outliers such that their impact will be minimized; these

[31] Dombois and Eckel (2011) divide the audification process into stages of data acquisition, signal condition and sound projection [27]. The methodology presented here focuses more directly on the use of audification as a tool for data analysis purposes.

methods include interpolation, value assignment, and exclusion. For the purposes of this guide, the term "interpolation" may encompass a wide variety of curve-fitting procedures (though nearest-neighbor interpolation should be considered a form of value assignment).[32]

Table 11. Various methods for handling "bad" or missing data values, reprinted from Alexander, et al., 2014.

	Flagged Data Values are _____.	Audible clicks and pops?	Preserves temporal progression of audio?
Interpolation	Smoothed	Minimized	Yes
Value Assignment	Assigned a constant value	Likely	Yes
Exclusion	Removed from Audio	Likely	No

6.1.3. **Data Scaling**

Data produced by scientific instruments may fall within a wide range of values, therefore most time series will need to be scaled within a standard range prior to the audio-translation process. The objective in this stage should be to maximize the signal-to-noise ratio while avoiding clipped data, as values that fall outside the amplitude range of the audio signal will be truncated, resulting in digital distortion. The appropriate range for data scaling may vary based on **1)** the bit depth of the audio, **2)** the file format, and **3)** the platform used for audification. The following equation is recommended for linearly scaling the data to a standard range of $-1 < x(t) < 1$.

$$x(t) = 2\frac{D(t) - \min D}{\max D - \min D} - 1 \qquad (2)$$

Where t denotes individual data samples, and $\min D$ and $\max D$ are the minimum and maximum data values in the array, respectively.

If a single data vector is split across multiple arrays, it may be preferable to concatenate the arrays or assign global values for $\min D$ and $\max D$. The use of global

[32] Some data sets may also require spike elimination prior to audification.

values in the creation of multiple audio files will result in uniform amplitude scaling, which is essential for comparative assessments across files.

6.1.4. **Audification**

The audification process involves the isomorphic (i.e., one-to-one) mapping of data samples to audio samples. Consequently, this process should not produce aliasing or artifacts, though auditory analysis may reveal artifacts that are inherent within the original data. The vast majority of modern waveform editors and media players are capable of importing and playing back audio files that span several hours in length. However, in most cases it will be preferable to write multiple files of shorter duration for ease of navigation and file sharing.

Several scientific computing languages such as Matlab and IDL contain built-in functions for writing audio waveforms from variables stored in memory.[33] Example routines for audification in Matlab and IDL are provided online in order to facilitate domain-specific research initiatives [208]. The primary parameters that need to be considered while creating an audio file are **bit depth**, **sampling rate**, the number of **audio channels**, and the **file format**.

6.1.4.1. **Bit Depth**

Bit depth refers to the number of bits used to encode the amplitude value of each sample in an audio file. While the industry standard for CD quality is 16 bits, 24 or 32 bits are recommended when the application of digital signal processing techniques is anticipated. **Table 12** provides the number of unique amplitude values that are resolvable through each associated bit depth, as well as the resulting dynamic range (the human auditory system has a functional dynamic range of approximately 120 dB [63]). So long as the bit depth of the audio file is greater than or equal to that of the original data, it may be possible to re-import a processed audio file into scientific computing software as a data vector, without any associated loss of signal resolution.

[33] These software toolkits also contain routines for playing back audio directly within the interface, though this functionality is generally limited.

Table 12. Values that may be useful in determining an appropriate bit-depth for the resulting audio file.

Number of bits	Unique Amplitude Values	Dynamic Range
16	65,536	96 dB
24	16,777,216	144 dB
32	4,294,967,296	385 dB

6.1.4.2. **Sampling Rate**

The sampling rate corresponds to the number of audio samples played back per second, measured in Hz. The industry standard rate (for CD quality playback) is 44.1 kHz, which enables the reproduction of spectral components up to 20.05 kHz (the maximum frequency that can be resolved within the audio file will occur at one-half the sampling rate, known as the Nyquist frequency [268]). Research has demonstrated that many trained listeners are able to perceive an increase in audio quality between signals played at 44.1 kHz and 88.1 kHz [269]. The human auditory system displays heightened sensitivity within the frequency range of 2–5 kHz, which should be taken into consideration while determining an appropriate sampling rate.[34] In the case of confirmatory analysis, it may be ideal to assign a sampling rate that will map the spectral components of target features within this spectral bandwidth (see equation 3 in section 6.1.8).

6.1.4.3. **Audio Channels**

In most instances it will be appropriate to audify each data parameter as a single-channel audio file, though distributing multiple parameters across audio channels will allow for spatialization effects (e.g., stereo panning). The localization of auditory streams may be exploited for the synchronous comparison of multiple data parameters; additionally, many software and hardware tools for monitoring audio loudness levels include correlation meters that may be visually assessed to determine correlation across data vectors (in real-time). Two channels (stereo) is the industry standard for playback over headphones and speakers, though many speaker systems support “5.1” and “7.1” surround-sound. In these systems, “5” and “7” refer to the number of speakers used for

[34] See Chapter II for a discussion of auditory sensitivity as a function of frequency.

spatialization (respectively), while ".1" refers to the low-frequency information sent to the subwoofer channel. There is no theoretical limit to the number of speakers that may be used for spatialization—wave field synthesis (WFS) techniques may utilize upwards of one-thousand speakers [270].

6.1.4.4. **File Format**

Uncompressed audio formats are strongly encouraged (e.g., WAVE and AIFF), as "lossy" encoding methods (e.g., MP3 compression) will remove frequency information from the file. One minute of audio at 16 bits / 44.1 kHz corresponds to a file size of approximately 10 MB.

6.1.5. **Repeated Close Listening for Feature Detection**

Once data have been written to an audio file format, the task of the analyst involves auditorily observing the frequency spectrum for unique patterns, features, and potentially anomalous events. Uncompressed audio can be imported and played within most modern media players. As a general precaution, the system volume should always be reduced before playing a new or unfamiliar audio file in order to avoid potential damage to the auditory system (which may result in permanent hearing loss). Routine listening breaks should be taken in order to allow the auditory system to "reset," as temporary threshold shifts can significantly alter the perception of spectral components (see Chapter II).

An inaccurate or low-quality reproduction of the audio signal (e.g., via laptop speakers) can introduce artificial distortion and resonances. High-fidelity speakers or headphones are recommended in order to maximize signal clarity during playback. Many headphones will inherently "color" an audio signal, and this often translates to an increased spectral power at low frequencies (manufacturers of most high-end equipment will supply the frequency profile as an equalization curve). Ideally, the audio signal should be reproduced with a relatively "flat" response across all frequencies, such that attention will not be biased toward (or away from) a given specific spectral bandwidth. While monitoring over speakers, phase cancellation and room-modes will reduce and enhance specific components within the frequency spectrum, and in the case of stereo

audio a “sweet spot” between the two speakers should be determined for maximum spatial resolution.

6.1.6. **Exploration and Manipulation**

Various software tools enable the real-time navigation and manipulation of digital audio waveforms (Audacity is a free cross-platform audio waveform editor). Certain software platforms, such as iZotope Rx, enable simultaneous auditory and visual displays (through spectrograms and spectral-power curves). Additionally, computer-music programming environments such as Max/MSP allow for the creation of custom interfaces for exploring audified waveforms.

6.1.6.1. **Navigating the Audified Waveform**

In the process of evaluating an audio file for features of importance, it may be helpful to first identify a well-known feature to establish a frame of reference. Repeated listening to any portion of an audified data set may reveal the presence of subtle underlying features. Many software toolsets contain functionality for placing markers within an audio file, which may be helpful for labeling and cataloguing regions in the data for later reference. The audified waveform may be regarded as a 1-dimensional line plot of the original time series, and can provide important visual feedback during the analysis process.

Many scientific data sets will contain a high level of broadband noise, as complex systems found in nature are often prone to turbulent behavior and non-linearity. In the case of the solar wind, the spectral slope of the audio file will be largely dependent on the time scale that was observed by the satellite instrument (see **Figure 47** for an overview of the turbulent solar wind spectrum).

6.1.6.2. **Digital Signal Processing**

Many digital signal processing (DSP) techniques are frequently applied within the domain sciences in the analysis of time series data, and a large number of these techniques are also commonly used by computer-music specialists to manipulate audio waveforms. For example, the application of high-pass, low-pass, band-pass, and notch filters may be useful for removing undesirable noise (e.g., spin-tone produced by a

satellite's rotation) or accentuating a portion of the frequency spectrum for closer auditory inspection. Noise reduction algorithms may also be applied in the investigation of underlying features, though these techniques should be applied with great care as they may produce significant spectral artifacts.

Increasing or decreasing the sampling rate (i.e., playback speed) of the audio will respectively shift spectral content to a higher or lower frequency space, revealing structures at larger and smaller scales, respectively. Algorithms that enable time expansion and compression (e.g., phase vocoding) are generally discouraged as they can produce a significant amount of spectral distortion.

Gain adjustments to a given subset of the audio file may help to balance the overall loudness, and normalization may be applied within most waveform editors to maximize the amplitude of the audio waveform. If individual clicks or pops are encountered, this range within the audio file may be attenuated to subjectively reduce their presence. Compression may be applied to audio signals with a large dynamic range in order to systematically attenuate loud regions (or amplify quiet regions).

6.1.7. **Recognition of Spectral Features**

Features of interest may occur at large and/or small time scales across a broad range of spectral power. As documented in Chapter II, the auditory system is particularly well suited for the task of separating meaningful signals from background noise. In the case of time series derived from satellite observations, equipment-induced artifacts will often sound "mechanical" while features of interest may assume a slightly more "organic" quality. It should become easier to quickly classify and categorize regions within the data as a variety of spectral features are observed over repeated listening sessions—in this way a "vocabulary of sounds" will expand over time. As spectral signatures in the audio are correlated with features in the data, it may be useful to adopt onomatopoeic naming conventions. While marking regions of interest within the audification, the use of both auditory and visual feedback may be helpful in defining precise temporal boundaries.

6.1.8. **Cross Referencing**

Sample numbers may be referenced within the audio file and mapped directly onto the source-range within the original data set so long as the one-to-one relationship has been preserved between data and audio samples (some DSP techniques will alter the temporal progression of the audio file). Frequency content can be mapped from the audio sampling rate to the cadence of the original time series by applying **equation 3**:

$$\boldsymbol{f_d = f_x \left(\frac{\mathbf{SR}_d}{\mathbf{SR}_a} \right)} \quad (3)$$

Where $\boldsymbol{f_d}$ is the frequency as found in the data, $\boldsymbol{f_x}$ is the frequency in the audio, $\mathbf{SR}_{\boldsymbol{d}}$ is the sampling rate of the data, and $\mathbf{SR}_{\boldsymbol{a}}$ is the sampling rate of the audio (all units measured in Hz). This equation may also be used to map spectral features observed in the original data set into the frequency space of the audio file.

6.1.9. **Confirmation Through Traditional Methods**

Once a region in the audio has been matched with a corresponding region in the original data, traditional data analysis methods may be applied to evaluate the time series. Spectral analysis techniques commonly applied within the heliospheric sciences include Wavelet and Fourier analysis. In many instances it may be necessary to investigate the identified time region across a range of available data parameters.

6.2. Conclusion

This chapter offered a framework for conducting auditory analysis through audification that may serve as a valuable reference for domain scientists who are interested in adopting this technique within their working practice. While this methodology was developed through interactions with space physicists, the analysis process was presented in such a way that knowledge should be readily transferable across scientific domains.

CHAPTER VII

Visual and Multimodal Identification of Wave-Particle Interactions in Heliospheric Time Series Data

Previous chapters highlighted the ability of the auditory modality to detect subtle periodicities and resonances within audified solar wind data, culminating in Chapter V where audification was used in the investigation of heliospheric time series data sets for both exploratory and confirmatory data analysis purposes. In the fourth and final case study, the audification specialist was able to expedite the analysis process of a research scientist by cataloguing the occurrence of LFWS events in *Wind* magnetometer data. This demonstrated a useful application for audification, and also leads to numerous questions including: **1)** Can domain scientists consistently and reliably apply the methods proposed in Chapter VI to generate audio files from solar wind time series? **2)** Can features of interest subsequently be identified with minimal training? **3)** How time-intensive is this process? This chapter investigates the above questions through a tightly controlled within-subjects study.

Seventeen scientists at NASA Goddard Space Flight Center (GSFC) and three members of the U of M Solar and Heliospheric Research Group (SHRG) were trained to audify solar wind time series and subsequently assess for LFWS events both visually and multimodally (visual analysis supplemented with auditory display). Supplementing visual analysis with auditory feedback resulted in the detection of 75% more events, though this multimodal analysis was approximately 2.9 times slower. The task results for all participants were compared against an expert assessment and it was found that while multimodal analysis displayed a higher sensitivity and overall balanced accuracy (measures applied from binary classification theory), it also led to a greater number of false-positive identifications. To better understand the nature of these false-positives, a

list of five regions selected by the majority of participants through multimodal analysis was provided to the expert for a follow-up assessment. The expert found that all five regions contained some level of Low Frequency Wave (LFW) activity, and though four of these regions did not meet LFWS criteria, one region contained an LFWS event that had not been identified in the original assessment. This supports the hypothesis that auditory analysis can be helpful in identifying features in time series data that may otherwise be missed through visual analysis methods.

7.1. Experimental Method

7.1.1. **Hypothesis**

The central hypothesis of this study is that by using multimodal analysis, participants will be able to more rapidly and successfully identify anomalous events that are characterized by temporal patterns unfolding in data streams derived from one-dimensional time series.

7.1.2. **Participants**

This study utilized a total of 20 participants: 17 researchers in the Heliophysics division at NASA GSFC and 3 members of the University of Michigan SHRG. The group included 15 males and 5 females age 22 to 71 (with an average age of 39), and 18 out of 20 had familiarity with spectrogram displays. With regard to educational experience, 13 participants held PhDs, 3 had completed a master's degree, and 4 held bachelor's degrees.

7.1.3. **Stimuli**

The examples used in this study were comprised of six time regions selected from data gathered by the magnetometer instrument on the *Wind* spacecraft from the year 2005. All regions were pre-screened to ensure they contained at least one region of clearly discernible wave-particle interactions. Datasets were downloaded from CDAWEB in ASCII format and imported into IDL for conversion to a ".sav" file format. The resulting files were assigned random numeric labels and provided to each participant in a single folder. Each dataset was approximately 2,646,000 samples in length, spanning 2

days 19 hours and 37 minutes of real time recording. The resulting audified data files were each approximately 1 minute in length. All examples were taken from the *z* component in the GSE coordinate system (which points north out of the plane of the earth's orbit around the sun). This component was chosen due to the relatively low presence of instrumentally induced noise produced by the rotation of the spacecraft (in comparison to the *x* and *y* components). These data samples were presented to participants in a randomized order. **Table 13** displays the precise time regions selected for each data Example.

Table 13 Data sources from the *Wind* MFI archive (*z*-component).

	START dd-mm-yyyy hh:mm:ss	**STOP** dd-mm-yyyy hh:mm:ss
Example 1	03-02-2005 14:27:46	06-02-2005 10:04:57
Example 2	03-03-2005 04:03:35	05-03-2005 23:40:46
Example 3	14-04-2005 07:45:27	17-04-2005 03:22:39
Example 4	25-04-2005 16:16:28	28-04-2005 11:53:40
Example 5	28-07-2005 13:58:02	31-07-2005 09:35:14
Example 6	23-11-2005 06:14:53	26-11-2005 01:52:05

7.1.4. **Apparatus**

The study was administered on a 15-in Macbook Pro with the Mac OS X 10.8.5 operating system. All instructions and training modules were provided in a web-based interface that incorporated several video-tutorials. The study and post-test were provided in Qualtrics, a web-based survey platform. Listening tasks were completed with Audio Technica ATH-M50 professional studio monitor headphones. Participants were instructed to apply a prompt-based script written in the IDL (v8.2.0) technical computing language to audify each dataset as a single-channel 16-bit audio file in the *.wav* file format with a sampling rate of 44,100 Hz. The analysis task was conducted using the iZotope Rx 3 software platform (version 2.10.656). All sessions took place in a quiet space that was free of distractions. While the majority of participants completed the entire

study in one session, a few participants required multiple sessions to complete the analysis task for both modalities.

7.1.5. **Procedure**

This study uses a within subjects design in which participants were asked to conduct an assessment of audified solar wind data sets within the iZotope Rx software package. Participants were trained to audify magnetometer data gathered by the *Wind* spacecraft and complete six analysis tasks presented in random order: **visual analysis** of three spectrogram displays, and **multimodal analysis** of three spectrogram displays *paired with audio playback*. The task involved the identification of LFWS activity by locating regions containing increased power within a specific spectral bandwidth. Participants were originally instructed to identify Ion Cyclotron Waves that were defined by their unique spectral signature. Participants proceeded to identify Ion Cyclotron Waves as well as Magnetosonic waves that produce nearly identical spectral signatures and vary only in polarization. The term "Low Frequency Wave Storm" was proposed by Jian et al. 2014 as it encapsulates waves with both right-handed and left-handed polarity. This terminology will be adopted for the remainder of this investigation as it more accurately encompasses the set of features identified by participants in this study. The training included several examples of LFWS events that had been previously confirmed by an expert (see **Figure 49**). The entire study took approximately 2 hours to complete on average (one hour for training and another for the analysis task). All participants indicated that they clearly understood the analysis task they were asked to perform, and were able to successfully identify a region of LFWS activity in a data example provided during a pre-test.

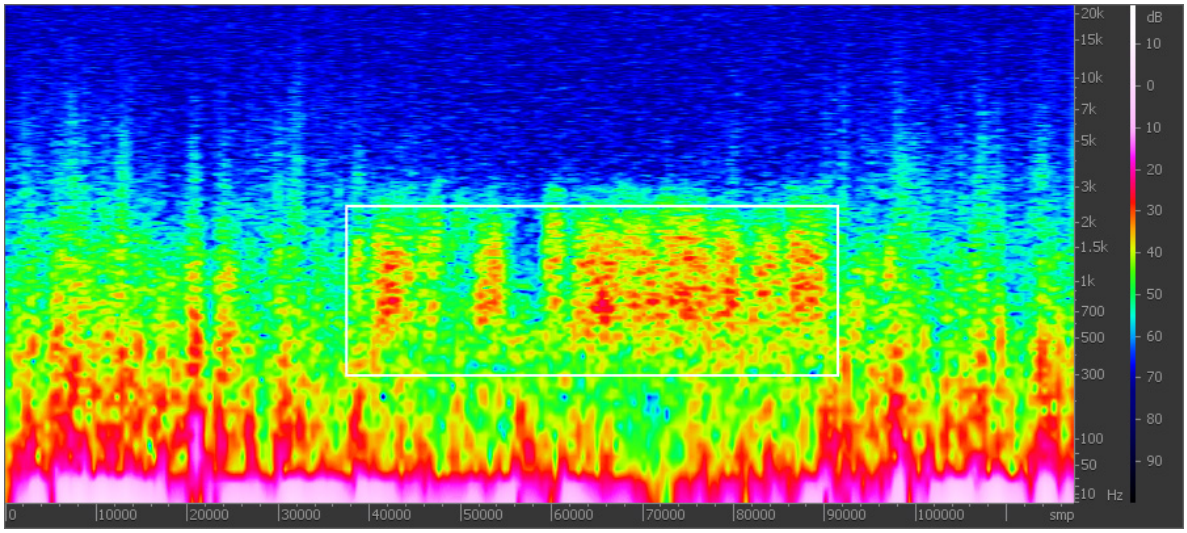


Figure 49. A region containing LFWS activity displayed in iZotope Rx (see white box). This example was used during the training session.

7.2. Results

In all instances, a two-tailed *t*-test was conducted to determine if the results were statistically significantly different from random chance performance. A matched *t*-test was used when comparing a participant's performance across the auditory and visual modalities, and an unmatched *t*-test was implemented when calculating performance within individual data examples (as no participant was presented with a single example in both modalities). Significance was considered at a value of $p < 0.05$, and strong significance at a value of $p < 0.01$.

As a metric for successful feature identification, participant assessment for the presence of LFWS events was compared against an analysis conducted by an expert (with several published papers on the topic of LFWS identification). The expert identified a total of 48 regions containing LFWS activity, with an average of 8 regions per file (the highest number in any given file was 14 and the lowest was 4). These regions varied in length from 6,687 data samples (0.15 seconds in the audification) to 50,683 data samples (1.15 seconds), with an average length of approximately 15,000 samples (0.34 seconds). In each case, a region was considered a positive match if the participant selected more than 50% of the data samples identified by the expert as containing a LFWS event. On average, participants correctly located 81% ± 8% of LFWS events through multimodal analysis and 46% ± 17% through visual analysis, indicating that participants correctly

identified approximately 78% more features when auditory display was available (p = 0.004). In total, multimodal analysis resulted in the positive identification of 176 more LFWS events than visual scan (an average of 8.8 features per participant). Multimodal analysis outperformed visual analysis across all 6 data examples, as displayed in **Figure 50**.

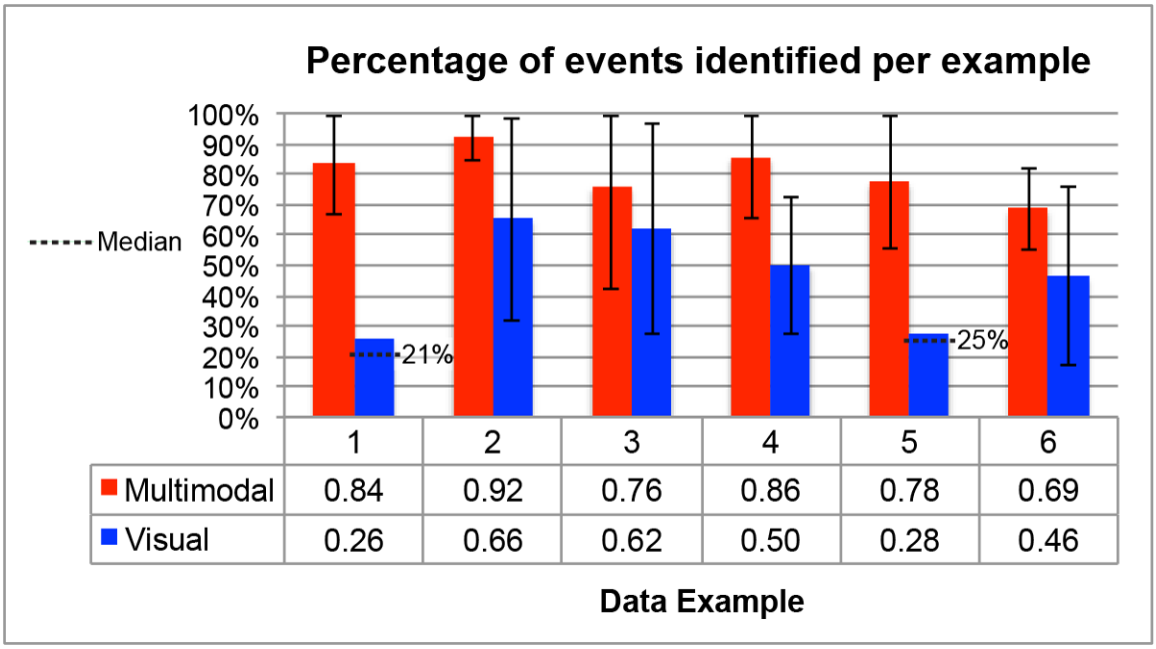


Figure 50. Multimodal analysis resulted in higher positive identification rates across all six data examples. Error bars indicate standard deviation. The median value is provided in lieu of error bars for instances in which the distribution is heavily skewed by an outlier (in these instances the small number of valid data points leads to a highly non-Gaussian distribution).

On average, participants took 42 ± 26 sec to generate audio files from the data; this was not included in the calculation of task completion time. Participants on average took 2.25 ± 0.99 min to complete each task using visual analysis, and 6.5 ± 0.73 min through multimodal analysis (p = 0.001), indicating that the multimodal search was roughly 2.9 times slower.

7.3. Discussion

7.3.1. **Interpreting the results**

The central hypothesis (that participants would be more successful in identifying LFWS events through the application of multimodal analysis) is quantitatively substantiated, as the application of multimodal analysis resulted in the positive identification of nearly twice as many events across all six data sets, and in the case of the first data example multimodal analysis uncovered more than three times the number of Low Frequency Wave Storms. These results were found to have strong statistical significance.

This study found a major disparity in completion time between visual and multimodal analysis tasks, with the visual task being completed roughly three times more quickly on average. The studies presented in Chapter III suggested that visual analysis may be faster than auditory/multimodal scan in some structured analysis tasks, though to a much smaller degree. Many participants in the study presented in this chapter had over a decade of experience working with scientific visualization, but few had any experience whatsoever with auditory analysis (as revealed through a post-test). For this reason this disparity in completion times could potentially be attributed to the novelty of the auditory analysis technique, as novelty effects are commonly reported in the literature of Human Computer Interaction design [271, 272]. A second study was designed to confirm the novelty effect through isolation of the auditory condition.[35]

7.3.2. **Investigating the Timing Disparity**

A follow-up study was conducted with six members of the SHRG to determine whether auditory analysis without any visual feedback would also result in a significantly longer completion time on the analysis task. This study emulated the design of the original experiment except for one condition—the visual analysis portion of the study was replaced with an auditory only task. It was hypothesized that if visual scan were inherently more expeditious, then pure auditory scan should be much slower than multimodal analysis, in which case visual scan was also made available.

[35] A longitudinal study of potential learning and novelty effects was deemed to be outside the scope of this dissertation, and is suggested as a valuable future study.

In total, participants in the follow-up study identified 76% ± 20% of features through multimodal analysis and 72% ± 19% of features through auditory analysis. The difference in performance between the two conditions was not found to be statistically significant, nor was the difference between either condition and the multimodal performance from the first study, hence, it was determined that auditory and multimodal identification rates were nearly equal and comparable to multimodal performance in the first study.

Participants in the follow-up study took an average of 5.3 ± 2.9 min to complete the multimodal analysis task, and 6.2 ± 3.2 min to complete the auditory analysis task. The difference in completion time between the two conditions was not found to be statistically significant, and while auditory analysis was slightly slower on average, both closely match the multimodal completion time from the initial study. These findings suggest that the novelty effect played a large role in the noted difference in completion time in the original study, and it is anticipated that extended training on this task would result in faster auditory/multimodal completion times.

7.3.3. **Binary Classification**

The original measure for successful task completion was the percentage of features identified by participants that were found to be in agreement with an expert assessment. This measure identifies the likelihood with which multimodal or visual analysis will result in the positive identification of a LFWS embedded within turbulent solar wind data. However, this provides a limited perspective of task performance, as results will always be positively biased toward the selection of larger regions (i.e., a participant who selects the entire duration of each audio file as containing LFWS activity will achieve a successful identification rate of 100%).

In order to achieve a more nuanced understanding of task performance, this investigation must not only consider the regions selected by the participants and the expert (true positive) but also the regions selected by participants that were not selected by the expert (false positive), regions selected by the expert that were not selected by the participant (false negative), and regions that were neither selected by the participant nor the expert (true negative). In treating each data sample in the audification as a yes/no

trial, standard measures from binary classification theory can be applied to determine participant **sensitivity**, **specificity**, and **balanced accuracy** (for positive "+" and negative "-" indication of LFWS activity) [273]:

		Expert +	Expert -
Participant	+	True Positive	False Positive
Participant	-	False Negative	True Negative

$$\textbf{Specificity} = \frac{\mathrm{TN}}{\mathrm{TN} + \mathrm{FP}} \qquad \textbf{Sensitivity} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}$$

$$\textbf{Balanced Accuracy} = \frac{\text{Sensitivity} + \text{Specificity}}{2} \tag{4}$$

In this study, **sensitivity** indicates the percentage of data samples that contain LFWS activity (as indicated by the expert) that were correctly identified by the participant, which provides the likelihood that a participant will correctly identify LFWS activity when it is present. Average participant sensitivity was 0.82 ± 0.11 through multimodal analysis, and 0.48 ± 0.27 through visual analysis (p = 0.0001),[36] indicating that participants were much more sensitive to the presence of LFWS events through multimodal analysis. This result closely mirrors that which was found using the original measure for successful feature identification rate (which is to be expected).

Specificity indicates the percentage of data samples that do not contain LFWS activity that were correctly identified by the participant, which provides insight as to the ability of participants to exclude the presence of LFWS activity when none is present. Average participant **s**pecificity was 0.70 ± 0.16 through multimodal analysis, and 0.79 ± 0.21 through visual analysis (p = 0.001), indicating that participants were better at excluding the presence of LFWS activity through purely visual analysis.

Just as sensitivity displays a bias toward the selection of larger regions, specificity is equally biased toward the selection of small regions, as the selection of no regions would result in perfect specificity. In order to reconcile these two values into a single measure, they may be weighed equally to assess overall **balanced accuracy** (see eq. 4). Overall balanced accuracy was 0.76 ± 0.08 through multimodal analysis and 0.64 ± 0.16 through visual analysis (p < 0.001), indicating that multimodal analysis yielded a more

[36] For all binary classification measures reported in this section, "1" indicates perfect performance while "0" indicates the worst possible performance (and likely a high level of confusion).

accurate identification of LFWS events in the identification task. The results for sensitivity, specificity and balanced accuracy are plotted in **Figure 51**.

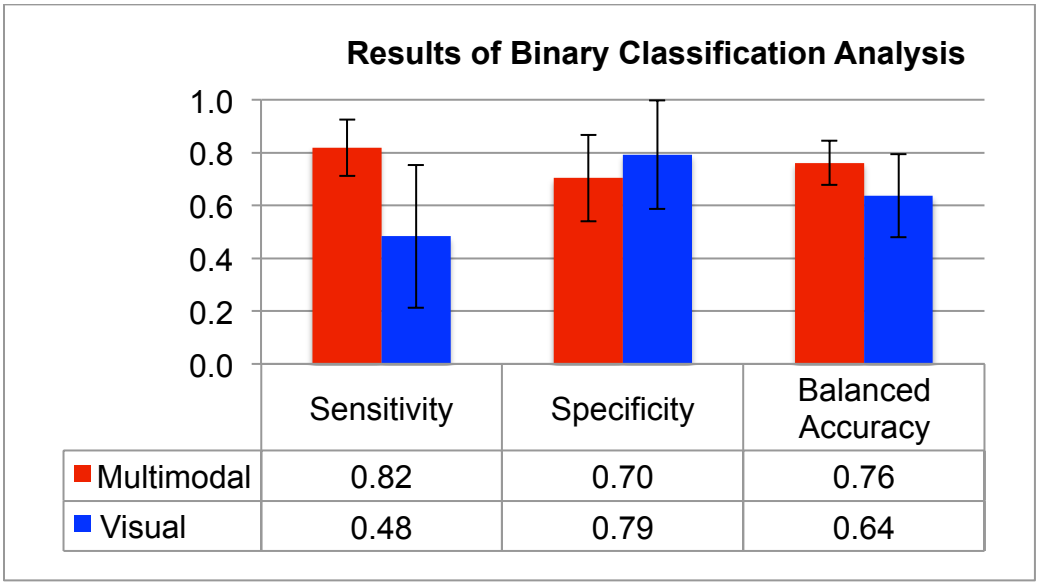


Figure 51. The results of the search task, compared against an expert assessment (a score of 1 indicates perfect performance).

At present, while many heliospheric scientists rely on visual scan to identify features of importance in spectrogram displays, participants in this study were 19% more accurate (considering both sensitivity and specificity) when the visual modality was supplemented with auditory feedback. Furthermore, the heightened sensitivity provided by multimodal analysis resulted in the positive identification of 76% more LFWS events across the 6 data examples. High sensitivity is important in many medical diagnostic tests where a false positive is less troublesome than a failure to diagnose a serious condition [274]. Heightened sensitivity is also useful while cataloguing features (such as LFWS events) in long time series, as the investigation and exclusion of false-positives is generally preferable to the creation and distribution of an incomplete catalogue. The research scientist in the fourth case study presented in Chapter V was presented with a catalogue of LFWS events in *Wind* data occurring during the year 2005, as identified through multimodal analysis. While this list contained a high percentage of false-positives, she estimated that her overall time expended in the analysis the 2005 archive was reduced by a factor of 10.

7.3.4. **Analysis of an Individual Data Example**

Binary classification measures were used to evaluate participant selections on a sample-by-sample basis across all six data examples, thus providing insight into the nature of regions identified across the two conditions. Yet these measures still do not tell the full story of this rich data set. A deeper level of insight can be gleaned from an examination of regions selected by the 20 participants within specific data examples. For this, it is helpful to visualize participant and expert selections alongside the original spectrogram display. **Figure 52** displays the spectrogram (reduced in size) for the second data example along with regions selected by the expert, regions selected by participants through multimodal and visual analysis, a single panel overlaying all participant selections, and a panel displaying solid coloration where 7/10 participant selections were in agreement. Heat maps for the remaining data examples are included in **Appendix II**.

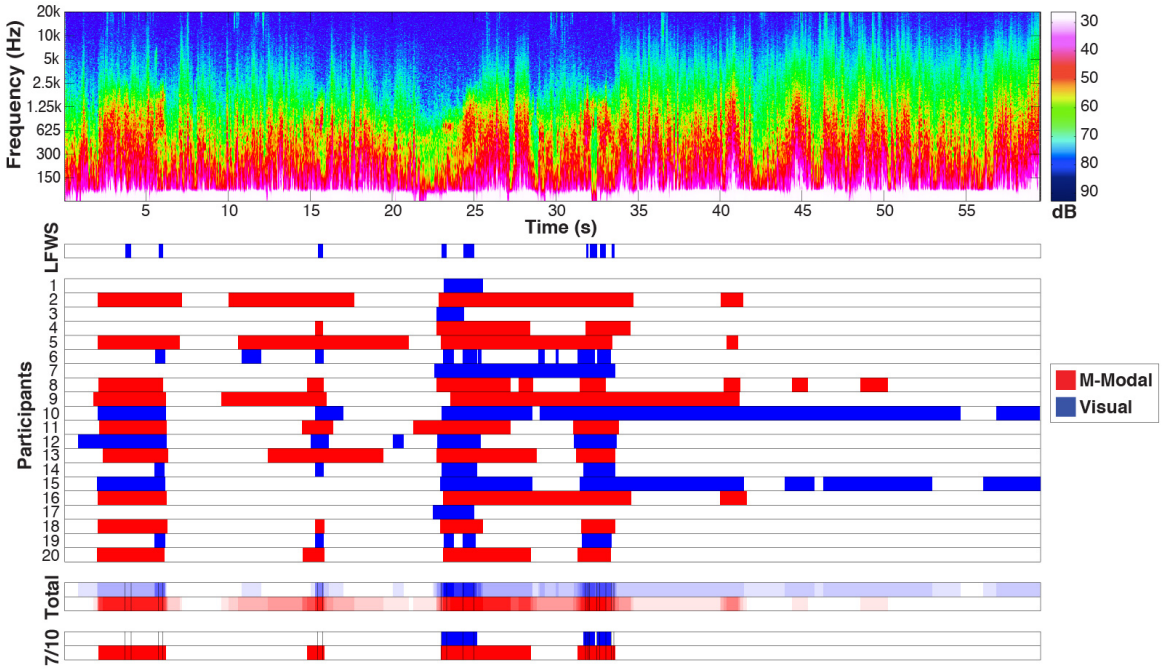


Figure 52. The results of the participant analysis for the second data example. The top panel contains a spectrogram representation of an audified data set from the *Wind* spacecraft. The next panel labeled "LFWS" indicates regions identified by the expert as containing LFWS activity. Below, highlighted areas indicate regions selected by each of the 20 participants through multimodal (red) and visual analysis (blue). The "Total" panel contains a summation of the selected regions across all participants, and the final panel contains solid coloration where 7/10 participants were in agreement.

Note that while the spectrogram display in **Figure 52** is provided as a static image, participants in the study were able to zoom in and out to examine potential features of interest more closely. From this macro-perspective, LFWS activity is a bit harder to pick out through visual inspection, but this can be done by noting the blue regions on the "LFWS" bar and exploring the corresponding range of the spectrogram. The LFWS activity can be seen as bright red regions of spectral enhancement between 300 Hz and 1.5 kHz, with prominent storm events occurring at 16 sec, 24 sec, and 33 sec into the file.

A quick visual scan of **Figure 52** reveals a wealth of information that was not immediately apparent through the results of the binary classification, namely that **1)** regions selected by the expert were generally very short in comparison to regions selected by participants, **2)** participant selections tended to aggregate in vertical bands around areas containing LFWS activity, **3)** in this specific data example 7/10 participants positively identified every instance of LFWS activity through multimodal analysis, and **4)** while most false-positive regions selected by participants were scattered, five participants using multimodal analysis agreed that one very small region contained LFWS activity, though this region was not identified by the expert (see 41 sec into the data example). An examination of all six data sets revealed these false-positives to be a common occurrence, sparking a deeper investigation into this final point.

7.3.5. **Investigating False Positives**

To better understand the nature of regions in which multimodal analysis resulted in a high concentration of false-positives across participants, a list containing five of these regions was provided to the expert for follow-up assessment. The expert found that all five regions contained some level of LFW activity, and the majority of these regions contained a higher concentration of LFW activity than is generally found in the turbulent solar wind. In most instances these regions did not meet LFWS classification because the background magnetic field was too variable, the wave packets were too short, or the signal to noise ratio was too high. However, one region (in the fourth data example) selected by 8/10 participants using multimodal analysis was found to contain a LFWS event that had been missed in the original expert assessment (see **Figure 53**).

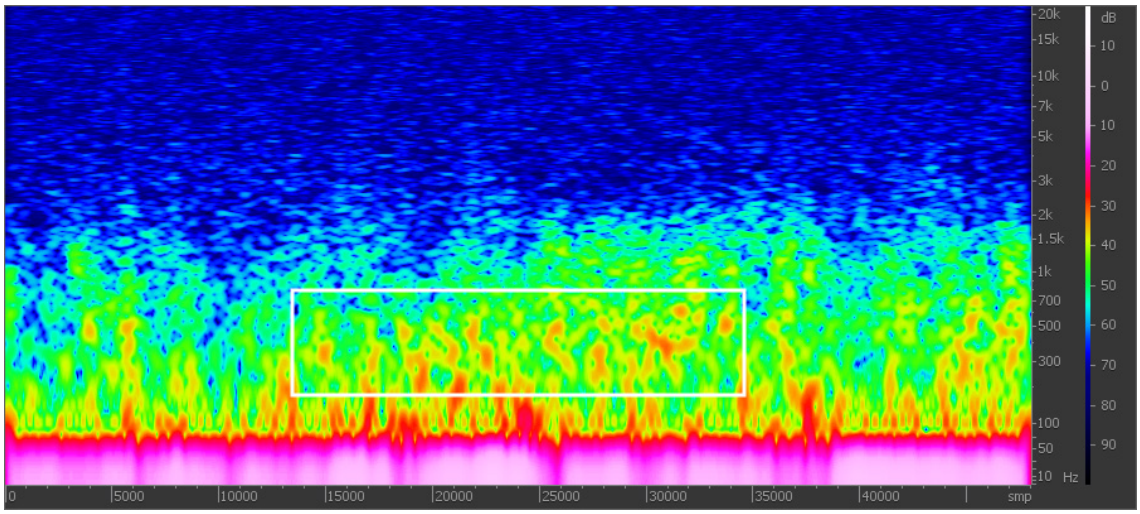


Figure 53. This region of LFWS activity (displayed here in iZotope Rx), was not identified in the initial expert assessment, but was identified by 8/10 participants through multimodal analysis (see white box).

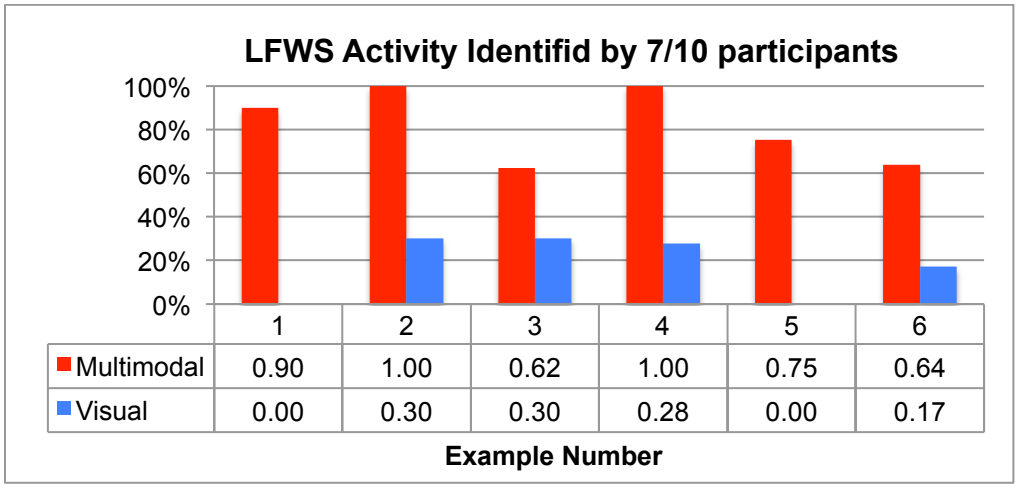
This LFWS event occurred on the 26th of April 2005 and spanned approximately 14,000 data samples (21.2 min in the data and 0.3 sec in the audification). While other instances of LFWS activity generally occurred in regions of the data where the turbulent background was less pronounced, this region was directly embedded in the presence of strong background turbulence. For comparison—**Figure 49** contains an example of LFWS activity that is isolated from the high-amplitude turbulence occurring along the bottom edge of the spectrogram display, this distinction is not immediately apparent in **Figure 53**. The boundaries of the LFWS event in **Figure 53** are more ambiguous, and the feature itself emerges as several isolated forms rather than one continuous visual object.

In the fixed-frequency identification study presented in Chapter III, artificial stimuli of varying length and amplitude were embedded within a masking signal [192]. Here it was found that auditory analysis outperformed visual analysis in the identification of extremely short events embedded within audified solar wind data sets, with the shortest events lasting 25 ms in length. Jian et al. (2014) defined LFWS events as persisting longer than 10 minutes in the data (translating to 6,600 data samples or 150 ms in the audification) and this criteria was not made known to the participants during the training session. However, the data contained a number of LFW events that were much shorter than 10 min—these features generally only last a few minutes, translating to

approximately 500–1,500 data samples or 10–35 ms in the audification. It is believed that participants likely identified a number of LFW events shorter than the 150 ms threshold for LFWS classification, leading to a large number of data samples flagged as false-positives.

7.3.6. **Performance Evaluated with an identification Threshold**

The final panel in **Figure 52** provides solid coloration in regions selected by seven out of ten participants—this threshold provides information in regard to both true-positive identification rates and the rate of agreement across participants. The specific value (seven out of ten) was selected because requiring agreement across less than half of participants was found to be too inclusive (resulting in the identification of broad regions of data), while requiring agreement of eight or more participants excluded almost all regions identified through visual assessment (resulting in very few instances of true positive identification). **Figure 54** displays the percentage of LFWS activity that is positively identified when this threshold is applied across all six data examples.



	1	2	3	4	5	6
Multimodal	0.90	1.00	0.62	1.00	0.75	0.64
Visual	0.00	0.30	0.30	0.28	0.00	0.17

Figure 54. Percentage of LFWS activity selected by 7/10 participants for each condition.

Using multimodal analysis, seven participants successfully identified 100% of LFWS activity through multimodal analysis in examples two and four, and the lowest identification rate was 62% for the third data example. Applying this same threshold to visual analysis—no LFWS activity was selected by at least seven participants in the first and fifth data examples, and the highest agreement percentage for any data example was

30% (in examples two and three). This result indicates that an average of 82% of the LFWS activity in each data set was identified by the majority of participants using multimodal analysis, while only 17% of LFWS activity was identified by the majority of participants through the application of visual analysis.

7.3.7. **Subjective Feedback**

A post-test questionnaire revealed that eighteen out of twenty participants found the auditory representation to be helpful in the analysis task, and fourteen reported a higher level of confidence in the task with the added auditory feedback. Seventeen participants indicated that they believed listening to data provided valuable scientific insight, and more than half reported that they would be interested in applying audification in the analysis of their own data in the future.

When asked to compare the experience of the visual and multimodal analysis tasks, several themes emerged. First, participants seemed to have a sense that multimodal analysis was more time intensive but resulted in a greater number of true-positive identifications—one participant described the visual scan to be "much quicker, but likely to miss subtle power," and another reported that multimodal analysis "required more time, but is more helpful." Second, participants noted that multimodal analysis provided a level of detail that was not accessible through purely visual scan, that it revealed "very fine structures in the data" and "made smaller features stand out." Third, several comments were made as to the qualitative contribution of the auditory feedback, and one participant noted a "characteristic wind/whistling sound that seems to be correlated to actual events." Lastly, many participants indicated that the auditory feedback increased their confidence in the decision making process, providing a way to "double-check" the identification of events. One participant stated that their decisions felt "less secure with no auditory input." These subjective reports support the early research of Hayward (1992), who (while studying the use of audification in geo-seismology) found that "auditory displays summarize the information in a different manner than common plots, allowing a second chance to recognize relationships in the seismic data" [16].

7.3.8. **Evaluation of Experimental Design**

At the outset of the study, the percentage of features successfully identified through each condition was established as the key metric for successful task completion. It was only after the data had been gathered that the decision was made to import the files into Matlab for a deeper analysis of sensitivity, specificity, and balanced accuracy. For this reason, these measures were not reported in the results section of the study; however, they do provide a level of detail far beyond the originally proposed metrics.

Several participants reported in a post-test survey that more and longer training examples would have been helpful in preparation for the analysis task. The training and pre-test included a total of 4 examples of LFWS activity (along with an example of unremarkable solar wind turbulence)—these training examples averaged 6.5 seconds in length, while each data example in the study spanned a full minute. The original rationale behind providing a limited number of pre-test data examples was to avoid over-training participants, however, it appears that the inclusion of at least one longer training example (30 sec or longer in length) would have helped to more adequately prepare participants for the identification tasks.

It was found that importing the raw data provided by CDAWeb into IDL took several minutes of processing time, regardless of available computational power. Hence the decision was made to provide the data to participants in the IDL *.sav* file format in order to reduce the importing time by several orders of magnitude. Any research scientist, regardless of whether they were importing data for audification or another form of traditional analysis, would encounter this bottleneck during the importing process. In most instances data would be audified directly from an IDL array already stored in memory, and it is this array that was provided to participants in the *.sav* file format.

A simple audification IDL code was provided to participants for practical purposes in order to reduce the likelihood of data loss due to a failure in the audification process. Several participants reported (in the post-test survey) that they believed it would have been more efficient to simply provide audio files rather than generating these files in real-time within IDL. This script provided a minimal set of controls similar to those that would be desired by a scientist working in the field, and hence it is believed that the audification process in this study mirrored the audio-file creation process that would be

encountered in the field, and the analysis task provided an effective valuation of the audification method established in Chapter VI.

7.3.9. **Data Integrity**

A post-test revealed one participant to be deaf in one ear with normal hearing in the other, while they were able to successfully complete the identification task, the data were excluded as there was no way to characterize how this hearing loss contributed to feature identification and selection. Another participant's data was excluded from the pool as they lacked sufficient experience with basic human-computer interactions and this drastically affected task performance, resulting in a malformed data set.

On several occasions participants failed to click the "next" button in the Qualtrics survey to indicate that they had completed the audification process and moved on to the analysis task. However, the total time for the audification and examination of each data file was still accessible and accurate in each case. In such instances, the time expended in the audification process was averaged for the remaining data examples, and this value was subtracted from the total time recorded for the example in question. This provided the most reasonable estimate for the duration of time spent in the analysis task. In each case, this value was checked against the completion time for the remaining analysis tasks in which time had been properly recorded, and in no case was this value found to be an outlier. This value was then used in place of the original (incorrect) timing data provided by the Qualtrics software.

The training module included several examples in which participants were instructed to select large regions of data containing LFWS activity, though these regions were identified by the expert as containing several sub-regions of activity with small gaps between events. In this way, participants were encouraged to err on the side of over-selection, which could have possibly contributed to an inflated number of false-positive identifications. This bias toward inclusion may have resulted in a lower overall accuracy, but it is not anticipated that this effect would be more prevalent in one condition or the other, and hence no preferential bias is anticipated.

It should be noted that the identification of LFWS events in turbulent solar wind data is a particularly difficult challenge regardless of the analysis method, and it is for

this reason that the task has eluded automation through statistical search procedures. While a balanced accuracy score below 0.9 would cast a failing grade on many medical diagnostic tests, a lower balanced accuracy is to be expected in this study as participants generally had no experience identifying LFWS events.

7.3.10. **Future Direction**

The work presented here captures the baseline identification abilities of participants that were provided minimal training. A future longitudinal study should track the abilities of participants over several sessions, spread out over the course of several weeks or months. This would provide insight into learning effects with regard to the relatively novel multimodal analysis task. It would also be a valuable opportunity for scientists to hone their multimodal analysis capabilities, as previous studies have shown that mere exposure to audified data sets increases the future identification of novel data sets [275].

Section 7.3.6 presented an analysis in which seven out of ten participants were required to agree on the presence of LFWS activity within a selected region in order for a positive identification to be made. This measure returns an extremely high identification rate when applied to regions selected through multimodal analysis, suggesting that this approach may be well suited for the type of distributed analysis commonly found in citizen science projects such as the highly successful *galaxy zoo* (a project that resulted in the visual classification of over one million galaxies by more than 100 thousand citizen volunteers) [276]. This study has demonstrated the effectiveness of web-based training modules, as well as the ability of participants to classify spectral phenomena in high-resolution solar wind data with very little training—laying the foundation for a citizen science project utilizing multimodal analysis techniques. Such a project would require a pipeline for distributing audified data, a front-end website for training volunteers and collecting data, and back-end coding to compile participant contributions, all of which are entirely feasible with minimal effort.

Code has been developed in coordination with the CDAWEB development team at NASA GSFC that would allow research scientists to download audified *Wind* MFI data directly from the web-repository. Providing access to a large archive of audified data on

such a widely accessible platform will certainly increase the visibility of this fledgling science and allow researchers around the world to capitalize on the strengths of this technique. It is not yet known when this functionality will become available.

7.4. Conclusion

The introduction to this chapter posed three questions asking whether domain scientists can reliably generate audio files from solar wind time series, identify features of interest with minimal training, and conduct this analysis in a timely fashion. The ensuing investigation demonstrated—across a pool of 20 researchers from NASA GSFC and the SHRG—that participants had no trouble using common heliospheric software tools to generate audio files from *Wind* MFI data. A large disparity in completion time was found between the two conditions, with participants taking approximately three times longer to complete the multimodal task. A follow-up study determined that this disparity is most likely attributable to novelty effects, though a longitudinal study will be required to test this hypothesis more rigorously.

Participants identified nearly twice as many LFWS events through multimodal analysis, and a deeper investigation applied measures from binary classification theory to determine that multimodal analysis also resulted in a higher overall balanced accuracy. Multimodal analysis displayed a high sensitivity, which contributed to a large number of false-positive identifications. Several regions that the majority of participants incorrectly identified as containing LFWS activity through multimodal analysis were provided to the expert for follow-up analysis, all of these regions were found to contain some level of LFW activity. Though four out of five of these regions did not contain sufficient activity to meet LFWS criteria, one region was found to contain a LFWS event that had originally been missed by the expert. This supports the previous assertion of Hayward (1994) that features may be identified through auditory means that would otherwise be missed through traditional analysis methods. The vast majority of participants found the auditory feedback to be helpful in the analysis task, and indicated that they would be interested in applying audification in future data analysis projects. This study provided crucial insight at the intersection of applied data analysis methods (in a real-world identification task) and thorough quantitative evaluation of participant performance.

CHAPTER VIII

Conclusion

8.1. Overview

The research presented in this dissertation investigated the use of audification as a method for the analysis of long time series data sets, with particular focus on the heliospheric sciences as a domain-specific case study. The scope of this investigation was defined by the three primary research questions established in Chapter I. First, what are the relative analysis capabilities of the visual and auditory modalities? A set of five empirical studies was conducted in which participants were sequentially tasked with the visual and auditory evaluation of time series from space instruments. A strong correlation was observed between assessments made across the two modalities, and auditory analysis demonstrated a significantly higher sensitivity to the presence of subtle spectral features. Second, can auditory analysis reveal spectral features missed through visual analysis methods? A number of open-ended analysis tasks revealed several instances in which features in the data were regularly missed through visual examination, and observed by the majority of participants through auditory analysis. Third, what potential benefits might audification provide when incorporated into traditional analysis workflows? A detailed case study was presented in which an experienced research scientist adopted audification methods into their regular working practice, and it was found that insights gleaned through auditory observation provided a valuable contribution to domain-specific research. Overall, this thesis demonstrates the value of audification for the analysis of large time-series, and substantiates the claim that auditory analysis can reveal knowledge that would otherwise be inaccessible through visual scan.

8.2. Primary Contributions

This dissertation provides three primary contributions to audification research: a comparative evaluation of auditory and visual analysis capabilities as provided by five empirical studies, a systematic method for the effective application of audification for exploratory data analysis, and a set of well-documented cases in which this method facilitated the acquisition of new knowledge in the heliospheric sciences. Chapter 1 provided a visual diagram that connected each research question with the area in the dissertation that directly addresses it. That figure is reprinted here for reference.

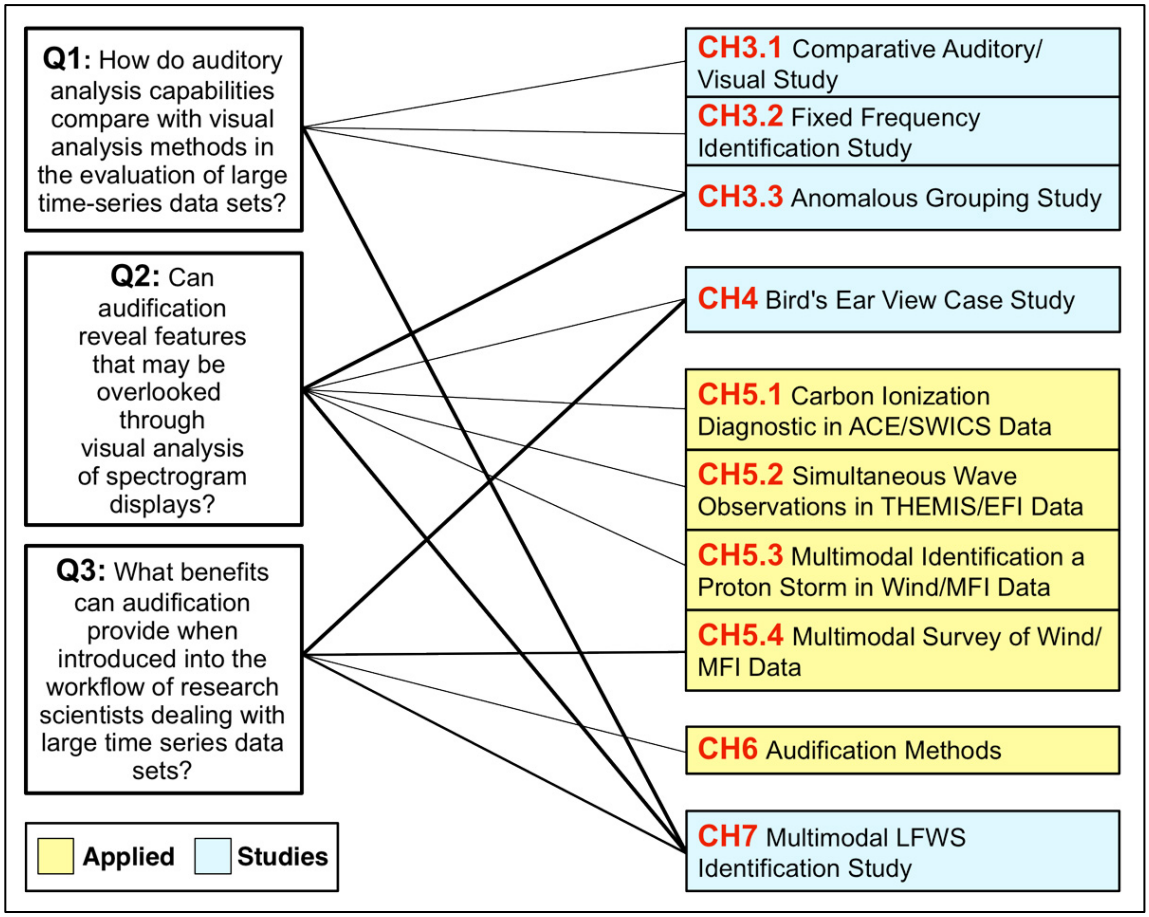


Figure 55. A map outlying the primary research questions and the chapter/section in which they are addressed. A line's thickness indicates the size of the contribution (reprinted from Chapter I, Figure 2).

The studies presented in Chapters IV and VII provide the primary quantitative and qualitative contribution of this research, while Chapters V and VI contain the primary contribution to audification methods and applications. The following section will provide

a detailed review of the larger contribution of this work framed within the context of each research question.

Q1: *How do auditory analysis capabilities compare with visual analysis methods in the evaluation of large time series data sets?*

The first study presented in Chapter III compared auditory and visual analysis techniques, establishing a strong correlation between assessments made through the two modalities in the evaluation of heliospheric data sets. A very strong correlation was found in the assessment of data gaps in audified data and spectrogram displays, and strong correlations were also found in the evaluation of the presence of broadband noise, repetitive elements, and fixed-frequency components. These findings were anecdotally supported in the case study presented in Chapter IV, where a research scientist was able to accurately describe the contents of a spectrogram display having only listened to the audification. The second study presented in Chapter III tasked participants with analyzing fixed-frequency sinusoids embedded in masking signals. Overall, visual analysis outperformed auditory analysis in the identification task, however, when performance was evaluated exclusively across data examples that contained embedded target-stimuli, the identification rate across the two modalities was approximately equal. Additionally, it was found that participants using visual analysis correctly identified a significantly higher percentage of sinusoids embedded within white-noise masks, while auditory analysis resulted in the correct identification of a significantly higher percentage of sinusoids embedded within audified data. Auditory analysis outperformed visual analysis in the identification of the shortest (25 ms) events, and true-positive identification rates loosely coincided with the frequency range of maximum auditory sensitivity.

In the final study, presented in Chapter VII, scientists were trained to audify time series gathered by the magnetometer on the *Wind* spacecraft and evaluate these for the presence of low-frequency wave storm (LFWS) events. Measures were incorporated from binary classification theory, known as **sensitivity** and **specificity,**[37] to provide a metric for participants' ability to correctly identify LFWS activity in regions where it was

[37] Sensitivity measures the percentage of data samples that were correctly identified as containing LFWS activity; specificity measures the percentage of data samples that were correctly identified as containing no LFWS activity.

present and to exclude activity in regions where it was not. Overall, multimodal analysis yielded a significantly higher sensitivity and lower specificity in this analysis task. Binary classification theory also suggested that these metrics could be combined, with equal weight, to provide a single measure of **balanced accuracy**. Overall, balanced accuracy was found to be significantly higher through multimodal analysis, in spite of the fact that this analysis method resulted in a much higher rate of false positive identifications. Additional investigation of these false-positives indicated that multimodal analysis was likely sensitive to the presence of low-frequency wave (LFW) events that were not long enough to meet the official LFWS criteria.

Regarding the efficiency of auditory analysis methods, all studies reported longer completion times for the auditory and multimodal conditions, though statistical significance was only found in two cases—the fixed-frequency analysis study in Chapter III, and the LFWS identification study presented in Chapter VII. The LFWS identification study presented an investigation into the observed disparity in task-completion times, and it was suggested that, as no participants had extensive experience working with auditory display, the difference may be largely attributable to the novelty effect.

Q2: *Can audification reveal features that may be overlooked through visual analysis of spectrogram displays?*

The fixed-frequency identification study presented in Chapter III investigated two specific instances in which features were missed by the majority of participants through visual analysis, but regularly identified auditorily. These examples were noted to contain sinusoids with an extremely short duration and low amplitude. The final study presented in Chapter III tasked participants with the identification of anomalous spectral features in several long time series gathered by the *Wind*/MFI instrument. Here, seven anomalous features were identified by the majority of participants through auditory analysis that were overlooked by all participants through visual analysis. The greatest disparity in feature identification rates was observed in the study presented in Chapter VII, in which participants correctly identified approximately 78% more LFWS events when visual display was supplemented with auditory feedback.

The work presented in this dissertation resulted in the auditory identification of four novel features in heliospheric time series data sets:

1) Instrumentally-induced noise in Ulysses magnetic-field data [277] that had yet to be documented in the literature (Chapter V) [205].
2) Spectral features in carbon ionic charge-states observed by the ACE/SWICS instrument that provide a new diagnostic for determining solar wind source regions (Chapter V.1) [222].
3) The simultaneous occurrence of wave modes in THEMIS/EFI data that had been predicted in theory but never observed in nature (Chapter V.2) [223].
4) An abnormally long and intense proton cyclotron wave storm in *Wind*/MFI data gathered during 2008 that served as the basis for an investigation of wave-particle interactions and mixing solar wind streams (Chapter V.3) [207].

While it cannot be said that these features would have been undetectable through alternative means, in several of these cases novel insight was produced in data archives that had been actively investigated within the heliospheric science community for over a decade. Additionally, these novel observations were generally produced quite rapidly; in Chapter V.2 the "chirp" sound generated by the simultaneous wave-modes was identified after several minutes of listening to audified THEMIS/EFI data. The detection of these novel features may be directly attributable to the high sensitivity of the auditory modality to the presence of subtle features embedded in complex spectra derived from solar wind time series, as documented in Chapter VII.

Q3: *What benefits can audification provide when introduced into the workflow of research scientists dealing with large time series data sets?*

Qualitative methods were applied to gain a sense for how this technique may directly contribute to longstanding (traditional) working practices within the domain sciences. Toward this end, the case study presented in Chapter IV incorporated a workflow analysis and applied the Think-Aloud protocol to gain access to the evolving thought processes of a research scientist as they applied audification toward feature identification and knowledge extraction—the scientist was able to successfully apply audification methods in the identification of subtle wave-particle interactions. In a follow-up interview, the research scientist stated that he continued to use audification for

the analysis of every large time series that he encountered, and it was noted that auditory analysis was particularly useful in the early stages of feature identification.

Hayward (1994) reported that geo-seismologists regularly listened to audio recordings of seismic data in order to locate earthquake events [16]—the case study in Chapter IV documented a similar application of audification by a space physicist in the exploration and evaluation of large heliospheric data archives. This case study also provided anecdotal support for the claim that audification is well suited for the identification of equipment-induced noise, as the scientist independently audified data from the Ulysses archive and successfully detected undocumented noise produced by the magnetometer instrument. The scientist also noted that the observance of small-scale features through auditory investigation provided a valuable new perspective on the data, to which he stated, "I'm going to change my whole method based on that knowledge." The subtle nuances unveiled through auditory display provided participants in the study presented in Chapter VII with a way to "double check" their work, and in the absence of auditory feedback their assessment felt "less secure." The sense of heightened confidence in task-performance conveyed in these subjective reports was quantitatively supported by the results of the analysis task.

The multimodal assessment of *Wind*/MFI data presented in the final section of Chapter V documents a case in which audification significantly contributed to a larger scientific investigation. Here the audification specialist applied auditory analysis methods to generate a catalogue of LFWS events in approximately 15 hours; the domain scientist estimated that this catalogue expedited her workflow by a factor of ten (reducing her time on task by approximately 72 hours). As many research scientists primarily rely on visual analysis for feature identification tasks, the adoption of auditory analysis methods could potentially bring about a corresponding increase in efficiency. To support this type of work, a set of methods and best practices is provided in Chapter VI.

8.3. General Discussion

Worrall (2010) recommended that the nature of features revealed through audification should be evaluated through controlled empirical studies—this dissertation picked up the baton and set about a systematic investigation guided by a clearly defined

set of research questions. This investigation fills a gap in the existing literature with respect to the baseline analysis capabilities of the auditory modality in comparison to traditional visual analysis methods. The acquisition of this foundational knowledge was a critical first step in the larger process of interaction design. Toward this end, this work drew from workflow observations to provide a systematic audification method that may be deployed across a wide range of hardware and software platforms. This method was applied within the heliospheric sciences to produce several domain-specific research outcomes, and was empirically tested in a study in which 20 heliospheric research scientists successfully evaluated solar wind time series through audio/visual analysis methods. Eighteen out of 20 participants found the auditory feedback to be helpful in completing the analysis task, and more than half showed interest in applying audification methods in their future research initiatives. The greatest hurdle to "gaining recognition of sonification as a valid research area" [1] may be exposing a larger set of researchers to auditory analysis methods such as those presented in this dissertation.

While the dissertation contributed several empirical studies to the growing body of audification literature, the nuanced spectral detail revealed through this analysis method has a quality that eludes quantification and verbal description. For this reason, a number of audio examples have been provided online [208], such that the reader may gain a sense for (what one research scientist described as) the "richness" of the information conveyed through this means—a subtle quality that Supper (2014) describes as "sublime."

The introduction provided a definition for data mining as a process for making large data sets "more understandable and useful." [5] Toward this end, the visual modality is well suited for a wide range of analysis tasks (to the extent that such a statement may be regarded as an inherent truism). However, this research suggests that auditory feedback may be preferable to visual analysis in many instances—and as a supplement to visual display, audification may reveal important features in large data sets that would otherwise be overlooked. In short, the valuable perspective provided by the *bird's ear view* is one that listening, alone, can provide.

8.4. Limitations

For the purposes of this dissertation the participant pool was limited to members of the SHRG and the heliospheric science division at NASA GSFC. A larger representative sample would increase the significance and statistical power of results in all cases.

A disparity was noted between completion times for auditory and visual analysis tasks. While the novelty effect was suggested, the extent of its role is uncertain. Additional longitudinal research would be necessary to determine the effect size. In all studies presented in this dissertation, participant pools were comprised of specialists, and hence results cannot be readily generalized to a larger population.

The anomalous-feature identification study presented in the final section of Chapter III noted a potential confounding impact of the presence of the visual waveform on the auditory analysis task. A subsequent study was conducted in which the waveform was removed that showed no impact on task performance.

The measures of sensitivity, specificity, and balanced accuracy (reported in the final study presented in Chapter VII) were adopted after the experiment had formally concluded, hence these outcomes were reported in the discussion section. The relative tendency of sensitivity to favor the selection of large regions within the data was reported, along with the tendency of specificity to favor the selection of small regions. It is not anticipated that the balanced accuracy measure would inherently produce an observable bias.

8.5. Future Work

Several longitudinal studies should be conducted to observe the impact of training on auditory and multimodal analysis tasks—such an investigation could also provide insight as to the effects of extended exposure on learning and novelty effects. Additionally, longitudinal case studies that incorporate data gathered in the field may result in the acquisition of new domain-specific knowledge. It is suggested that, if possible, audio and video recording should be integrated for thorough documentation, as it is impossible to predict when such outcomes will occur. The Think-Aloud technique [206, 209] may be helpful in developing new interfaces for applying audification in

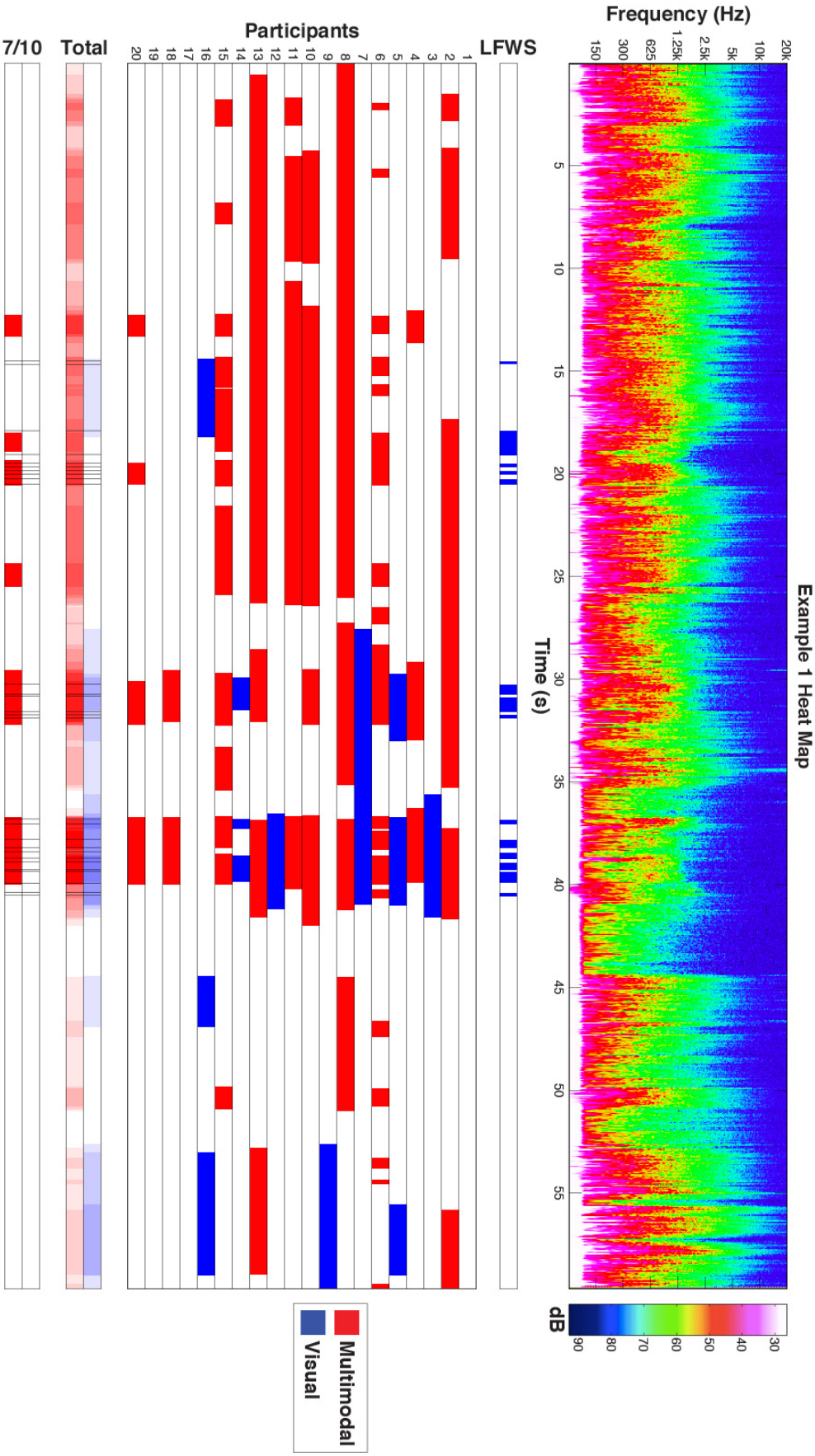
domain-specific contexts, and the Wizard-of-Oz paradigm [279] may also prove to be helpful in an iterative design process.

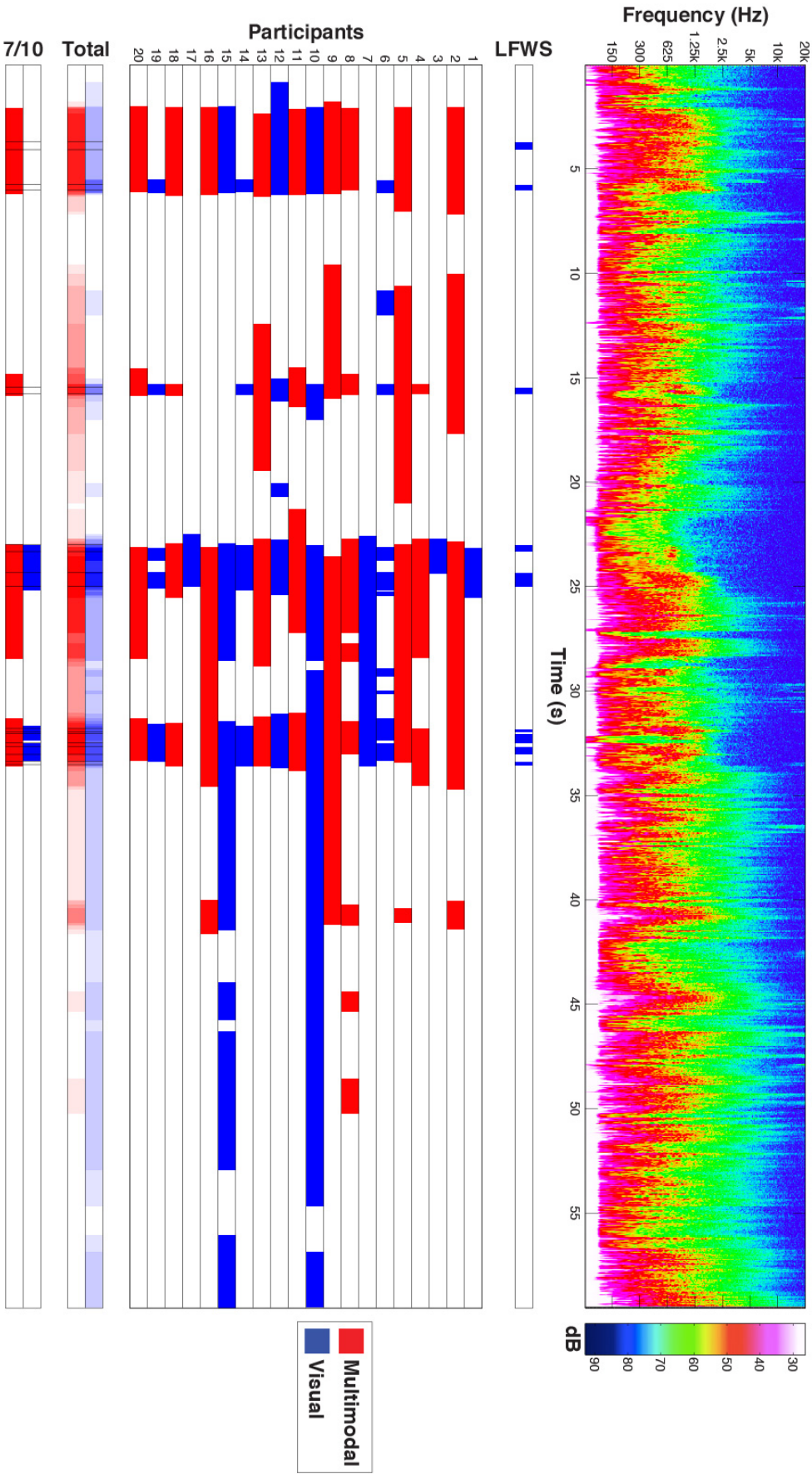
Future research should take advantage of public interest in this technique to crowd source the audio-visual analysis of large time series. An audification code has been provided to the CDAWeb team at NASA GSFC, with the ultimate goal of incorporating this web-repository into such a project. The increased accessibility of audified data through such an interface will provide an excellent resource for space scientists, the sonification community, and the public at large.

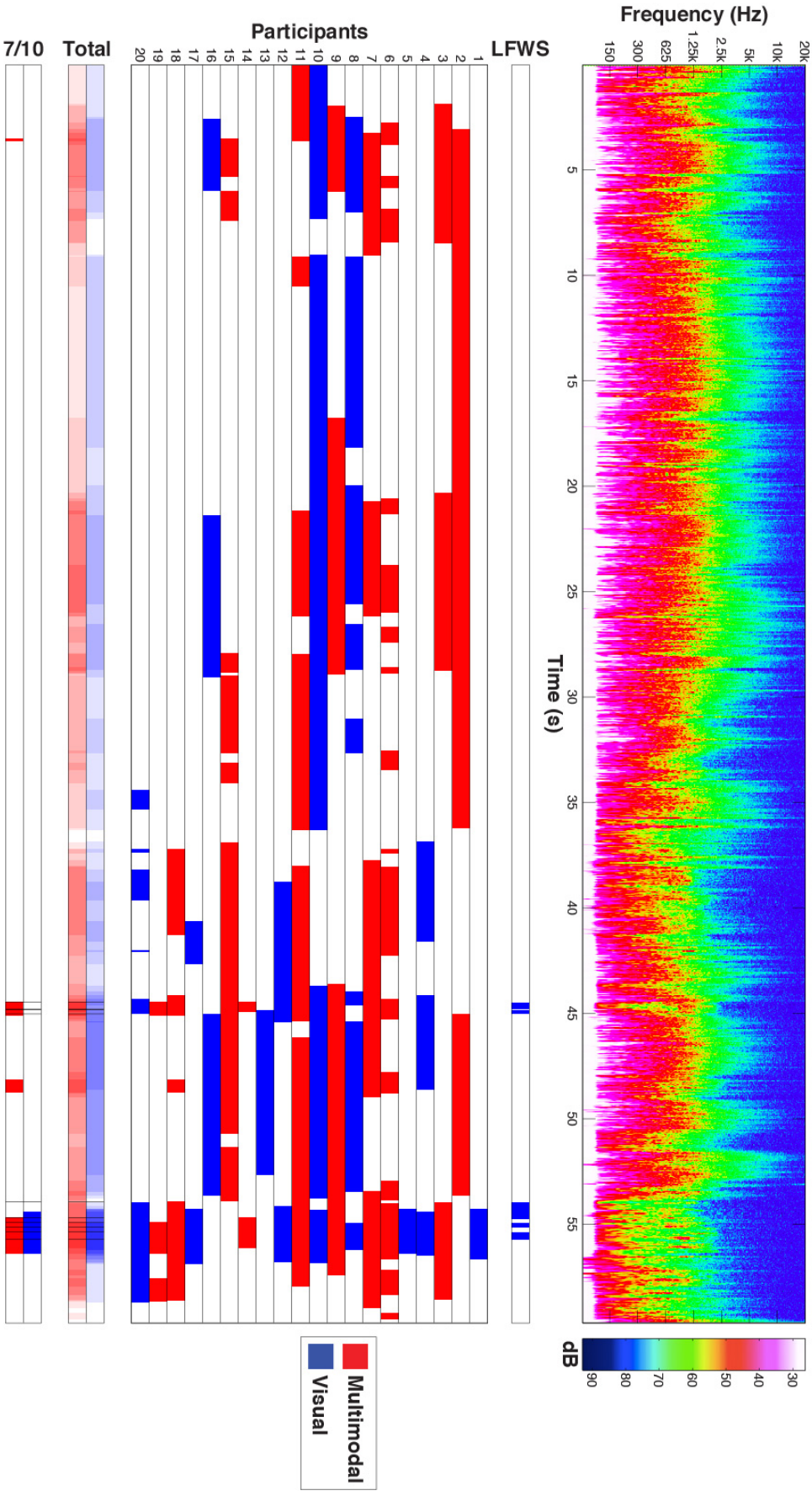
Appendix I – List of data examples used in the study presented in section 1 of Chapter III.

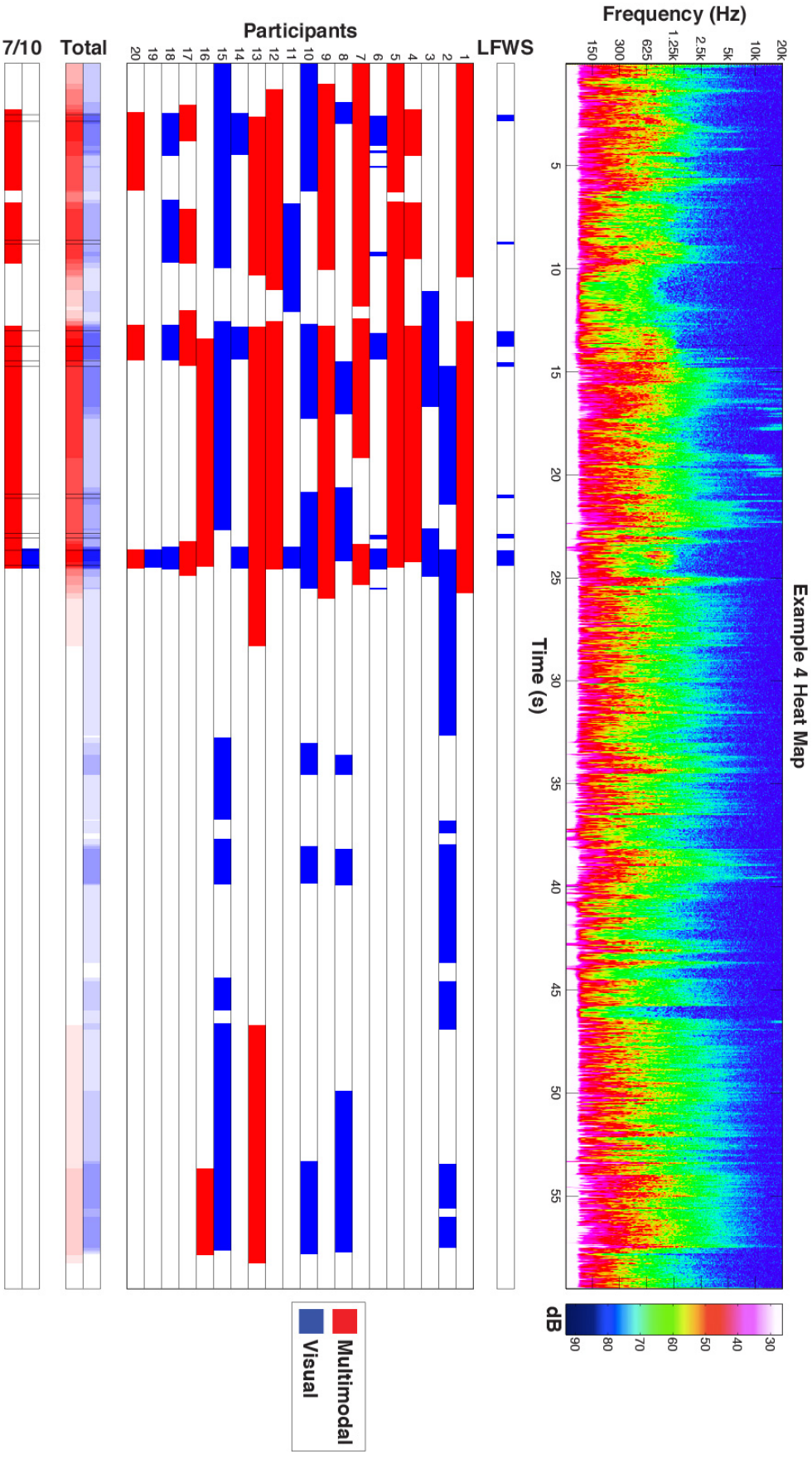
Example	Description
1	Magnetometer data from the ACE spacecraft. Contains a frequency component at 3.7 kHz
2	Magnetometer data from the ACE spacecraft. Contains strong repetitive bursts of broadband noise caused by rotating solar magnetic field lines.
3	OMNI plasma speed measured from 1963 to 1970. Contains large data gaps. Same data file as example 9.
4	Artificial data generated with Logic Pro. Contains repetitive broadband noise and a sinusoidal waveform at 500Hz.
5	Artificial data generated with Logic Pro. Contains broadband noise, data gaps, low frequency harmonic content, and a sinusoidal waveform at 2 kHz.
6	Artificial data generated with Logic Pro. Contains broadband noise, and low frequency harmonic content.
7	Artificial data generated with Logic Pro. Emulates the rise and fall of energy caused by a single shock, as occasionally found in solar wind data.
8	Data gathered from the FIPS instrument on the MESSENGER spacecraft. Contains strong repetitive elements from Mercury magnetopause crossings.
9	OMNI plasma speed measured from 1963 to 1970. Contains large data gaps. Same data file as example 3.
10	Helium velocity data from the SWICS instrument on the ACE spacecraft (1998-2009). Contains low frequency harmonic components.
11	Magnetometer data from the *Wind* satellite. The first half of the data file is repeated such that the file contains a single large repetitive element.
12	An audio recording of terrestrial wind. Contains resonances at approximately 300 Hz and 700 Hz. Data gaps were inserted into the recording.

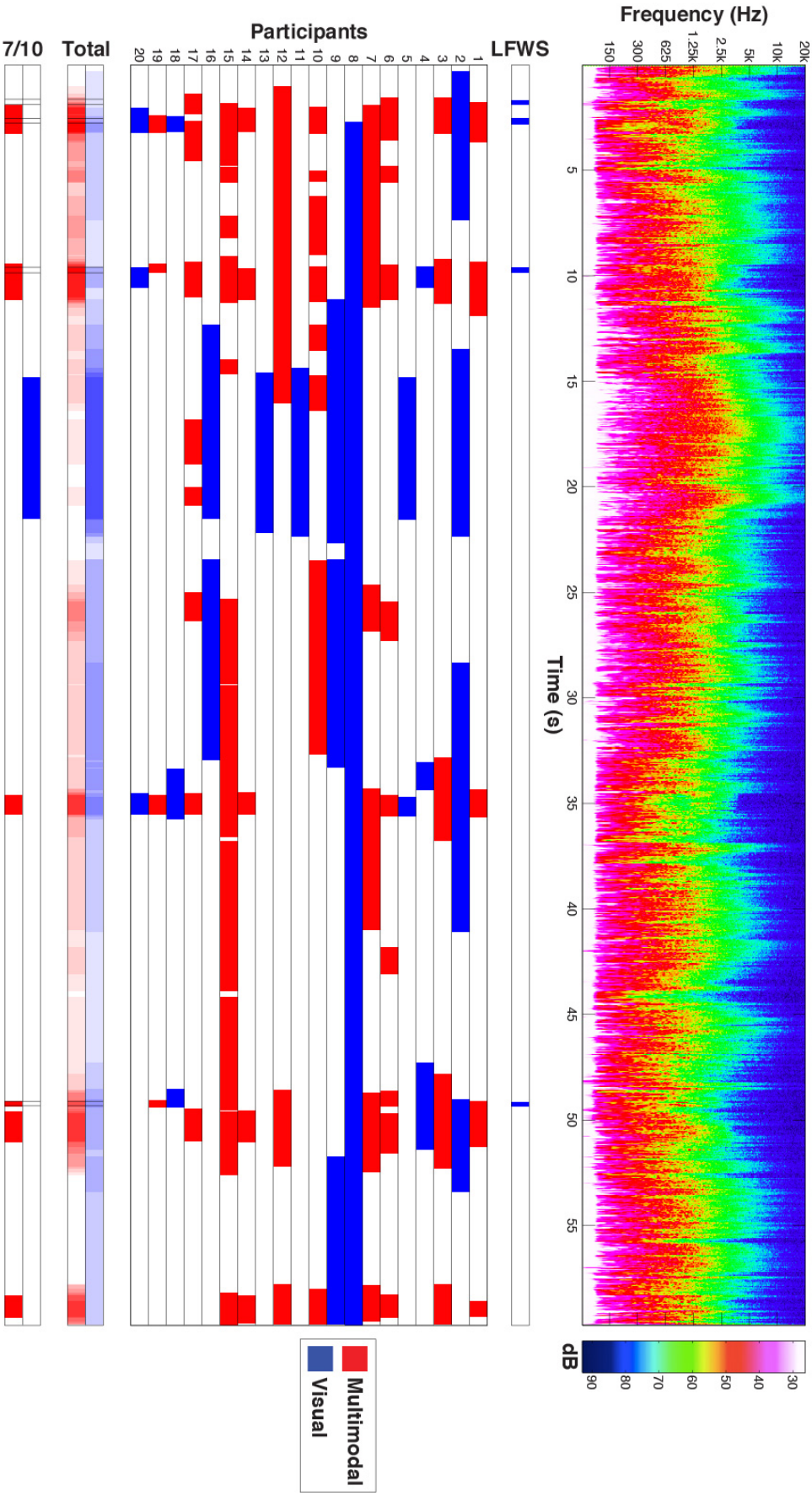
Appendix II – All heat maps for the LFWS study (Chapter VII).

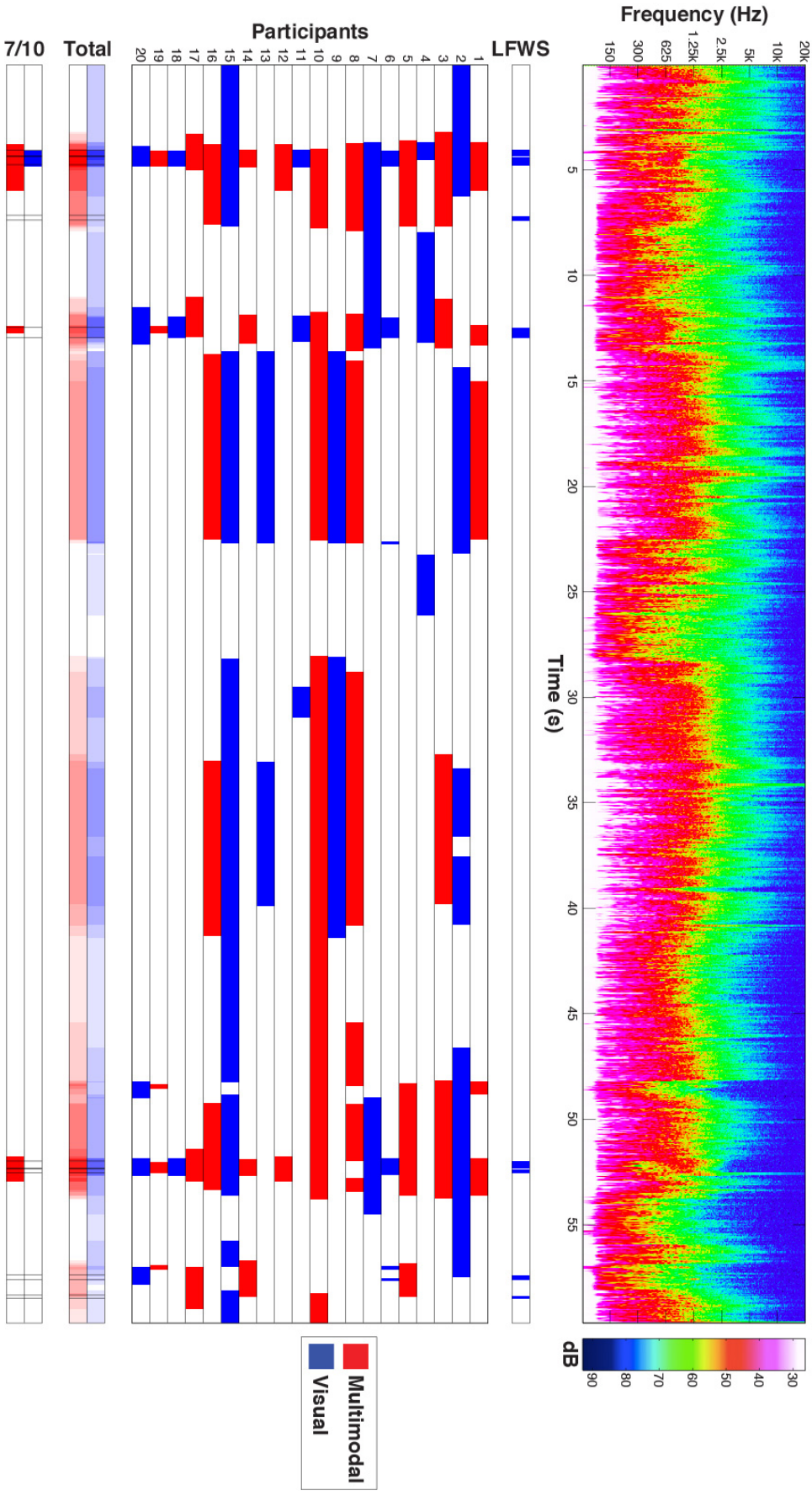



Example 2 Heat Map
Frequency (Hz)
20k
10k
5k
2.5k
1.25k
625
300
150
Time (s)
5
10
15
20
25
30
35
40
45
50
55
dB
90
80
70
60
50
40
30
LFWS
Participants
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
Total
7/10
Multimodal
Visual


Example 3 Heat Map
Frequency (Hz)
20k
10k
5k
2.5k
1.25k
625
300
150
Time (s)
5
10
15
20
25
30
35
40
45
50
55
dB
90
80
70
60
50
40
30
LFWS
Participants
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
Total
7/10
Multimodal
Visual


Example 4 Heat Map
Frequency (Hz)
20k
10k
5k
2.5k
1.25k
625
300
150
Time (s)
5
10
15
20
25
30
35
40
45
50
55
dB
90
80
70
60
50
40
30
Participants
LFWS
Total
7/10
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
Multimodal
Visual


Example 5 Heat Map
Frequency (Hz)
20k
10k
5k
2.5k
1.25k
625
300
150
Time (s)
5
10
15
20
25
30
35
40
45
50
55
dB
90
80
70
60
50
40
30
LFWS
Participants
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
Total
7/10
Multimodal
Visual


Example 6 Heat Map
Frequency (Hz)
20k
10k
5k
2.5k
1.25k
625
300
150
dB
90
80
70
60
50
40
30
Time (s)
5
10
15
20
25
30
35
40
45
50
55
Participants
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
LFWS
Total
7/10
Multimodal
Visual

REFERENCES

[1] G. Kramer, B. Walker, T. Bonebright, P. Cook, J. Flowers, N. Miner*, et al.*, *Sonification report: status of the field and research agenda*: International Community for Auditory Display, 1999.

[2] M. M. Gaber, A. Zaslavsky, and S. Krishnaswamy, "Data Stream Mining," in *Data Mining and Knowledge Discovery Handbook*, O. Maimon and L. Rokach, Eds., ed: Springer US, 2010, pp. 759-787.

[3] F. G. Hill, J.; Martens, P.; Bogart, R.; Davey, A.; Hourcle, J.; Suarez Sola, F.; Hughitt, K.; Spencer, J.; Reardon, K.; Amezcua, A., "SDO Data Access Via The Virtual Solar Observatory," in *American Astronomical Society, AAS Meeting #216, #402.18;* , 2010, p. 876.

[4] K. P. Reardon and S. Berukoff, "Improved data exploitation for DKIST high-resolution solar observations," in *Big Data (Big Data), 2014 IEEE International Conference on*, 2014, pp. 45-52.

[5] D. J. Hand, H. Mannila, and P. Smyth, *Principles of data mining*: MIT press, 2001.

[6] P. Cabena, P. Hadjinian, R. Stadler, J. Verhees, and A. Zanasi, *Discovering data mining: from concept to implementation*: Prentice-Hall, Inc., 1998.

[7] D. T. Larose, *Discovering knowledge in data: an introduction to data mining*: John Wiley & Sons, 2014.

[8] M. Berry and G. Linoff, *Mastering data mining: The art and science of customer relationship management*: John Wiley & Sons, Inc., 1999.

[9] J. W. Tukey, *Exploratory Data Analysis*. Reading, Massachusetts: Addison-Wesley Publishing Company, 1977.

[10] I. Pollack and L. Ficks, *Information of Elementary Multidimensional Auditory Displays* vol. 26: ASA, 1954.

[11] S. Fidell, *Sensory Function in Multimodal Signal Detection* vol. 47: ASA, 1970.

[12] N. E. Loveless, J. Brebner, and P. Hamilton, "Bisensory presentation of information," *Psychological Bulletin,* vol. 73, pp. 161-199, 1970.

[13] E. S. Yeung, "Pattern recognition by audio representation of multivariate analytical data," *Analytical Chemistry,* vol. 52, pp. 1120-1123, 1980/06/01 1980.

[14] S. Bly, "Presenting information in sound," presented at the Proceedings of the 1982 conference on Human factors in computing systems, Gaithersburg, Maryland, United States, 1982.

[15] T. Hermann, A. Hunt, and J. G. Neuhoff, *The sonification handbook*: Logos Verlag, 2011.

[16] C. Hayward, "Listening to the Earth Sing," in *Auditory display: Sonification, Audification, and Auditory Interfaces*, G. Kramer, Ed., ed Reading, Mass.: Addison-Wesley, 1994, pp. 369-404.

[17] F. Dombois, "Using Audification in Planetary Seismology," in *International Conference on Auditory Display*, Espoo, Finland, 2001, pp. 227-230.

[18] F. Dombois, "Auditory Seismology, On Free Oscillations, Focal Mechanisms, Explosions, and Synthetic Seismograms," in *International Conference on Auditory Display*, Kyoto, Japan, 2002, pp. 1-4.

[19] S. Pauletto and A. Hunt, "A Comparison of Audio and visual Analysis of Complex Time-Series Data Sets," in *International Conference on Auditory Display*, Limerick, Ireland, 2005, pp. 175-181.

[20] D. Worrall, "Using Sound to Identify Correlations in Market Data." vol. 5954, S. Ystad, M. Aramaki, R. Kronland-Martinet, and K. Jensen, Eds., ed: Springer Berlin / Heidelberg, 2010, pp. 202-218.

[21] W. L. Diaz-Merced, R. M. Candey, N. Brickhouse, M. Schneps, J. C. Mannone, S. Brewster*, et al.*, "Sonification of Astronomical Data," *Proceedings of the International Astronomical Union,* vol. 7, pp. 133-136, 2011.

[22] F. L. Scarf, D. A. Gurnett, W. S. Kurth, and R. L. Pynter, "Plasma Observations Near Saturn: Initial Results from Voyager 2," *Science,* vol. 215, pp. 563-570, January 29, 1982 1982.

[23] R. A. Helliwell, *Whistlers and related ionospheric phenomena*. Stanford, California: Stanford University Press, 1965.

[24] S. Sazhin and M. Hayakawa, "Magnetospheric chorus emissions: A review," *Planetary and space science,* vol. 40, pp. 681-697, 1992.

[25] J. LaBelle and R. Treumann, "Auroral Radio Emissions, 1. Hisses, Roars, and Bursts," *Space Science Reviews,* vol. 101, pp. 295-440, 2002/08/01 2002.

[26] E. J. Smith and B. T. Tsurutani, "Magnetosheath lion roars," *Journal of Geophysical Research,* vol. 81, pp. 2261-2266, 1976.

[27] F. Dombois and G. Eckel, "Audification," in *The Sonification Handbook*, A. H. Thomas Hermann, John G. Neuhoff, Ed., ed Berlin GmbH: COST 2011, p. 564.

[28] L. P. Feigen, "Physical characteristics of sound and hearing," *The American journal of cardiology,* vol. 28, pp. 130-133, 1971.

[29] T. D. Rossing, F. R. Moore, and P. A. Wheeler, *The Science of Sound: Third Edition*: Addison-Wesley Reading, MA, 2002.

[30] M. Koch, "The neurobiology of startle," *Progress in neurobiology,* vol. 59, pp. 107-128, 1999.

[31] A. T. Welford, "Introduction: an historical background sketch," in *Reaction Times*, J. T. Brebner and A. T. Welford, Eds., ed New York: Academic Press, 1980.

[32] J. Shelton and G. P. Kumar, "Comparison between auditory and visual simple reaction times," *Neuroscience & Medicine,* vol. 1, p. 30, 2010.

[33] E. Borg, "A quantitative study of the effect of the acoustic stapedius reflex on sound transmission through the middle ear of man," *Acta oto-laryngologica,* vol. 66, pp. 461-472, 1968.
[34] G. Von Békésy and E. G. Wever, *Experiments in hearing* vol. 8: McGraw-Hill New York, 1960.
[35] S. P. Bacon, R. R. Fay, and A. N. Popper, *Compression: from cochlea to cochlear implants*: Springer, 2004.
[36] B. C. J. Moore, *An introduction to the psychology of hearing (5rd ed.)*. Great Britain: Elsevier Academic Press, 2004.
[37] A. Wright, A. Davis, G. Bredberg, L. Ulehlova, and H. Spencer, "Hair cell distributions in the normal human cochlea," *Acta oto-laryngologica. Supplementum,* vol. 444, pp. 1-48, 1986.
[38] S. Carlile, "Psychoacoustics," in *The Sonification Handbook*, A. H. Thomas Hermann, John G. Neuhoff, Ed., ed Berlin GmbH: COST 2011, p. 564.
[39] M. Mathews, "The auditory brain," in *Music, cognition, and computerized sound*, R. C. Perry, Ed., ed: MIT Press, 1999, pp. 11-20.
[40] J. L. Lauter, P. Herscovitch, C. Formby, and M. E. Raichle, "Tonotopic organization in human auditory cortex revealed by positron emission tomography," *Hearing research,* vol. 20, pp. 199-205, 1985.
[41] A. American Standards and A. Acoustical Society of, *American tentative standards for sound level meters for measurement of noise and other sounds, approved American Standards Association, February 17, 1936, sponsor Acoustical Society of America, American Engineering and Industrial Standards*. New York, N.Y.: American Standards Association, 1936.
[42] P. M. Rabinowitz, "Noise-induced hearing loss," *American Family Physician,* vol. 61, pp. 2759-2760, 2000.
[43] S. Stevens, "The quantification of sensation," *Daedalus,* vol. 88, pp. 606-621, 1959.
[44] R. L. S. Pierre Jr, D. J. Maguire, and C. S. Automotive, "The impact of a-weighting sound pressure level measurements during the evaluation of noise exposure," in *Conference NOISE-CON*, 2004, pp. 12-14.
[45] H. Fletcher and W. A. Munson, "Loudness, its definition, measurement and calculation," *Journal of the Acoustical Society of America,* vol. 5, pp. 82-108, 1933.
[46] E. Zwicker, G. Flottorp, and S. S. Stevens, "Critical band width in loudness summation," *The Journal of the Acoustical Society of America,* vol. 29, pp. 548-557, 1957.
[47] S. S. Stevens, J. Volkmann, and E. B. Newman, "A scale for the measurement of the psychological magnitude pitch," *Journal of the Acoustical Society of America,* vol. 8, pp. 185-190, 1937.
[48] E. Zwicker and H. Fastl, *Psychoacoustics: facts and models*. Berlin: Springer Verlag, 1990.
[49] M. M. Van Wanrooij and A. J. Van Opstal, "Contribution of head shadow and pinna cues to chronic monaural sound localization," *The Journal of neuroscience,* vol. 24, pp. 4163-4171, 2004.
[50] I. J. Hirsh, "Masking of speech and auditory localization," *International Journal of Audiology,* vol. 10, pp. 110-114, 1971.
[51] H. Wallach, E. B. Newman, and M. R. Rosenzweig, "A Precedence Effect in Sound Localization," *The Journal of the Acoustical Society of America,* vol. 21, pp. 468-468, 1949.
[52] E. Shaw, "Earcanal pressure generated by a free sound field," *The Journal of the Acoustical Society of America,* vol. 39, pp. 465-470, 1966.
[53] S. Carlile, *Virtual auditory space: Generation and applications*: RG Landes Austin TX, USA, 1996.
[54] A. M. Mayer, "LXI. Researches in acoustics," *Philosophical Magazine Series 5,* vol. 2, pp. 500-507, 1876.
[55] R. Wegel and C. Lane, "The auditory masking of one pure tone by another and its probable relation to the dynamics of the inner ear," *Physical Review,* vol. 23, p. 266, 1924.
[56] A. S. Association, "Acoustic terminology," *New York: American Standards Association, S,* vol. 1, 1960.
[57] H. Fastl and E. Zwicker, *Psychoacoustics: facts and models*, Third Edition ed. Heidelberg: Springer, 2007.
[58] H. Fletcher, "Auditory patterns," *Reviews of modern physics,* vol. 12, p. 47, 1940.
[59] J. Hood, "Fundamentals of Identification of Sensori-neural Hearing Loss," *British Journal of Audiology,* vol. 6, pp. 21-26, 1972.
[60] R. J. Salvi, R. P. Hamernik, and D. Henderson, "Response patterns of auditory nerve fibers during temporary threshold shift," *Hearing Research,* vol. 10, pp. 37-67, 1983.
[61] J. Corey, *Audio Production and Critical Listening: Technical Ear Training*: Elsevier/Focal Press, 2010.
[62] H. L. Helmholtz, *On the Sensations of Tone as a Physiological Basis for the Theory of Music*: Cambridge University Press, 2009.
[63] C. Roads, *The computer music tutorial*: MIT press, 1996.
[64] A. S. Bregman, "The formation of auditory streams," *Attention and performance,* vol. 7, pp. 63-76, 1978.
[65] A. S. Bregman, *Auditory Scene Analysis: The Perceptual Organization of Sound*. Cambridge, Massachusetts: The MIT Press, 1990.
[66] A. S. Bregman and S. Pinker, "Auditory streaming and the building of timbre," *Canadian Journal of Psychology/Revue canadienne de psychologie,* vol. 32, p. 19, 1978.
[67] E. C. Cherry, "Some experiments on the recognition of speech, with one and with two ears," *The Journal of the acoustical society of America,* vol. 25, pp. 975-979, 1953.
[68] G. Fant, *Acoustic theory of speech production: with calculations based on X-ray studies of Russian articulations* vol. 2: Walter de Gruyter, 1971.

[69] J. M. Chowning, "The Synthesis of Complex Audio Spectra by Means of Frequency Modulation," *Computer Music Journal,* vol. 1, pp. 46-54, 1977.
[70] J. Chowning, "Perceptual fusion and auditory perspective," in *Music, cognition, and computerized sound*, 1999, pp. 261-275.
[71] G. A. Miller and G. A. Heise, "The trill threshold," *The Journal of the Acoustical Society of America,* vol. 22, pp. 637-638, 1950.
[72] I. Rock, *Indirect perception*: Mit Press, 1997.
[73] W. Köhler, "Gestalt psychology," 1929.
[74] K. Koffka, *Principles of Gestalt psychology*: Routledge, 2013.
[75] A. S. Bregman and P. A. Ahad, *Demonstrations of auditory scene analysis: The perceptual organization of sound*: Auditory Perception Laboratory, Psychology Department, McGill University, 1995.
[76] A. Bregman. (2008, Jan 25). *Auditory Scene Analysis*. Available: http://webpages.mcgill.ca/staff/Group2/abregm1/web/downloadstoc.htm - 01
[77] A. S. Bregman and J. Campbell, "Primary auditory stream segregation and perception of order in rapid sequences of tones," *Journal of experimental psychology,* vol. 89, p. 244, 1971.
[78] A. Bidet-Caulet and O. Bertrand, "Neurophysiological mechanisms involved in auditory perceptual organization," *Frontiers in neuroscience,* vol. 3, p. 182, 2009.
[79] W. T. Greenough, J. E. Black, and C. S. Wallace, "Experience and brain development," *Child development,* pp. 539-559, 1987.
[80] K. Wagner and K. R. Dobkins, "Synaesthetic associations decrease during infancy," *Psychological science,* vol. 22, pp. 1067-1072, 2011.
[81] D. Lewkowicz and G. Turkewitz, "Cross-modal equivalence in early infancy: Auditory--visual intensity matching," *Developmental Psychology,* vol. 16, pp. 597-607, // 1980.
[82] P. Walker, J. G. Bremner, U. Mason, J. Spring, K. Mattock, A. Slater*, et al.*, "Preverbal infants' sensitivity to synaesthetic cross-modality correspondences," *Psychological Science,* vol. 21, pp. 21-25, 2010.
[83] B. Bond and S. Stevens, "Cross-modality matching of brightness to loudness by 5-year-olds," *Attention, Perception, & Psychophysics,* vol. 6, pp. 337-339, 1969.
[84] L. E. Marks, R. Szczesiul, and P. Ohlott, "On the cross-modal perception of intensity," *Journal of Experimental Psychology: Human Perception and Performance,* vol. 12, p. 517, 1986.
[85] I. P. Howard and W. B. Templeton, "Human spatial orientation," 1966.
[86] G. J. Thomas, "Experimental study of the influence of vision on sound localization," *Journal of Experimental Psychology,* vol. 28, p. 163, 1941.
[87] L. Marks, E. Ben-Artzi, and S. Lakatos, "Cross-modal interactions in auditory and visual discrimination," *International journal of psychophysiology: official journal of the International Organization of Psychophysiology,* vol. 50, p. 125, 2003.
[88] W. Teramoto, S. Hidaka, and Y. Sugita, "Sounds move a static visual object," *PLoS One,* vol. 5, p. e12255, 2010.
[89] I. H. Bernstein and B. A. Edelstein, "Effects of some variations in auditory input upon visual choice reaction time," *Journal of Experimental Psychology,* vol. 87, p. 241, 1971.
[90] E. Ben-Artzi and L. E. Marks, "Visual-auditory interaction in speeded classification: Role of stimulus difference," *Attention, Perception, & Psychophysics,* vol. 57, pp. 1151-1162, 1995.
[91] S. B.-C. If, L. Burtlf, F. Smith-Laittan, J. Harrison, and P. Bolton, "Synaesthesia: prevalence and familiality," *Perception,* vol. 25, pp. 1073-1079, 1996.
[92] A. O'Leary, Rhodes, Gillian, "Cross-modal effects on visual and auditory object perception," *Perception & Psychophysics,* vol. 35, pp. 565-569, 1984.
[93] R. Sekuler, A. B. Sekuler, and R. Lau, "Sound alters visual motion perception," *Nature,* vol. 385, p. 308, 1997.
[94] M. O. Ernst, "A Bayesian view on multimodal cue integration," *Perception of the human body from the inside out,* pp. 105-131, 2005.
[95] A. d. Campo, "A Data Sonification Design Map," in *International Workshop on Interactive Sonification*, York, UK, 2007, pp. 1-4.
[96] B. N. Walker and M. A. Nees, "An agenda for research and development of multimodal graphs," in *International Conference on Auditory Display (ICAD2005), Limerick, Ireland*, 2005.
[97] S. Bly, "Sound and computer information presentation," UCRL-53282; Other: ON: DE82015782 United StatesOther: ON: DE82015782Wed Feb 06 18:33:31 EST 2008NTIS, PC A06/MF A01.LLNL; ERA-07-048959; EDB-82-135933English, 1982.
[98] S. P. Frysinger, *Applied research in auditory data representation* vol. 1259: SPIE, 1990.
[99] C. Scaletti and A. B. Craig, "Using Sound to Extract Meaning From Complex Data," in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*. vol. 1459, E. J. Farrell, Ed., ed, 1991, pp. 207-219.
[100] G. Kramer, Ed., *Auditory Display: Sonification, Audification and Auditory Interfaces,* (SFI Studies in the Sciences of Complexity. Addison-Wesley Publishing Company, 1994, p.^pp. Pages.

[101] W. W. Gaver, "Auditory icons: using sound in computer interfaces," *Hum.-Comput. Interact.,* vol. 2, pp. 167-177, 1986.
[102] W. W. Gaver, "The SonicFinder: an interface that uses auditory icons," *Hum.-Comput. Interact.,* vol. 4, pp. 67-94, 1989.
[103] W. W. Gaver, R. B. Smith, and T. O'Shea, "Effective sounds in complex systems: the ARKOLA simulation," presented at the Proceedings of the SIGCHI conference on Human factors in computing systems: Reaching through technology, New Orleans, Louisiana, United States, 1991.
[104] M. M. Blattner, D. A. Sumikawa, and R. M. Greenberg, "Earcons and icons: their structure and common design principles," *Hum.-Comput. Interact.,* vol. 4, pp. 11-44, 1989.
[105] E. Brazil and M. Fernström, "Auditory icons," *The sonification handbook,* pp. 325-338, 2011.
[106] D. McGookin and S. Brewster, "Earcons," 2011.
[107] J. J. Mezrich, S. Frysinger, and R. Slivjanovski, "Dynamic Representation of Multivariate Time Series Data," *Journal of the American Statistical Association,* vol. 79, pp. 34-40, 1984.
[108] D. L. Mansur, M. M. Blattner, and K. I. Joy, "Sound graphs: A numerical data analysis method for the blind," *Journal of Medical Systems,* vol. 9, pp. 163-174, 1985.
[109] T. L. Bonebright, M. A. Nees, T. T. Connerley, and G. R. McCain, "Testing the effectiveness of sonified graphs for education: A programmatic research project," in *Proc. Int. Conf. Auditory Display*, 2001, pp. 62-66.
[110] A. D. Edwards, "Auditory display in assistive technology," *The Sonification Handbook,* pp. 431-453, 2011.
[111] T. Hermann, "Model-based sonification," *The Sonification Handbook,* pp. 399-427, 2011.
[112] T. Hermann and H. J. Ritter, "Listen to your Data: Model-Based Sonification for Data Analysis," presented at the Advances in Intelligent computing and multimedia systems, Baden-Baden, Germany, 1999.
[113] T. Hermann and H. Ritter, "Neural gas sonification-Growing adaptive interfaces for interacting with data," in *Information Visualisation, 2004. IV 2004. Proceedings. Eighth International Conference on*, 2004, pp. 871-878.
[114] T. Bovermann, T. Hermann, and H. Ritter, "Tangible data scanning sonification model," in *ICAD Proceedings*, 2006, pp. 77-82.
[115] T. Hermann and H. Ritter, "Crystallization sonification of high-dimensional datasets," *ACM Transactions on Applied Perception (TAP),* vol. 2, pp. 550-558, 2005.
[116] K. Vogt, W. Plessas, A. De Campo, C. Frauenberger, G. Eckel, and D. A. Graz, "SONIFICATION OF SPIN MODELS. LISTENING TO PHASE TRANSITIONS IN THE ISING AND POTTS MODEL," in *Proc. of the ICAD, 13th International Conference on Auditory Display, Montreal, Canada*, 2007.
[117] T. Hermann, M. H. Hansen, and H. Ritter, "Sonification of Markov Chain Monte Carlo Simulations," in *International Conference on Auditory Display*, Espoo, Finland, 2001, pp. 1-9.
[118] F. Grond and J. Berger, "Parameter mapping sonification," *The Sonification Handbook,* pp. 363-397, 2011.
[119] G. Dubus and R. Bresin, "A Systematic Review of Mapping Strategies for the Sonification of Physical Quantities," *PloS one,* vol. 8, p. e82491, 2013.
[120] P. Vickers, "Sonification for process monitoring," in *The Sonification Handbook*, A. H. Thomas Hermann, John G. Neuhoff, Ed., ed Berlin GmbH: COST 2011, pp. 455-492.
[121] A. Kovaric, "New methods for counting the alpha and the beta particles," *Physical Review,* vol. 9, pp. 567-568, 1917.
[122] F. Kilander and P. Lönnqvist, "A whisper in the woods-an ambient soundscape for peripheral awareness of remote processes," 2002.
[123] E. Brazil and M. Fernstrom, "Investigating ambient auditory information systems," 2007.
[124] W. T. Fitch and G. Kramer, "Sonifying the body electric: Superiority of an auditory over a visual display in a complex, multivariate system," in *Auditory Display. Sonification, Audification, and Auditory Interfaces*. vol. 18, G. Kramer, Ed., ed Reading, MA: Addison-Wesley, 1994, pp. 307-325.
[125] M. Watson and P. Sanderson, "Sonification Supports Eyes-Free Respiratory Monitoring and Task Time-Sharing," *Human Factors: The Journal of the Human Factors and Ergonomics Society,* vol. 46, pp. 497-517, Fall 2004 2004.
[126] G. Baier, T. Hermann, and U. Stephani, "Event-based sonification of EEG rhythms in real time," *Clinical Neurophysiology,* vol. 118, pp. 1377-1386, 2007.
[127] G. Baier and T. Hermann, "Sonification: listen to brain activity, Music that works," R. Haas and V. Brandes, Eds., ed: Springer Vienna, 2009, pp. 11-23.
[128] G. Baier, T. Hermann, and U. Stephani. (2007). *Sonification Examples for: Multi-Channel Sonification of Human EEG*. Available: http://www.techfak.uni-bielefeld.de/ags/ami/datason/demo/ICAD2007/EEGSon/index.html
[129] M. W. Rodger, W. R. Young, and C. M. Craig, "Synthesis of walking sounds for alleviating gait disturbances in Parkinson's disease," *Neural Systems and Rehabilitation Engineering, IEEE Transactions on,* vol. 22, pp. 543-548, 2014.
[130] R. Sigrist, G. Rauter, R. Riener, and P. Wolf, "Augmented visual, auditory, haptic, and multimodal feedback in motor learning: A review," *Psychonomic bulletin & review,* vol. 20, pp. 21-53, 2013.

[131] R. Kirby, "Development of a real‐time performance measurement and feedback system for alpine skiers," *Sports Technology,* vol. 2, pp. 43-52, 2009.
[132] M. Takahata, K. Shiraki, Y. Sakane, and Y. Takebayashi, "Sound feedback for powerful karate training," in *Proceedings of the 2004 conference on New interfaces for musical expression*, 2004, pp. 13-18.
[133] J. Stienstra, K. Overbeeke, and S. Wensveen, "Embodying complexity through movement sonification: case study on empowering the speed-skater," in *Proceedings of the 9th ACM SIGCHI Italian Chapter International Conference on Computer-Human Interaction: Facing Complexity*, 2011, pp. 39-44.
[134] A. Godbout and J. E. Boyd, "Corrective sonic feedback for speed skating: a case study," in *Proceedings of the 16th international conference on auditory display*, 2010, pp. 23-30.
[135] N. Schaffert, K. Mattes, and A. O. Effenberg, "Listen to the boat motion: acoustic information for elite rowers," in *Human Interaction with Auditory Displays–Proceedings of the Interactive Sonification Workshop*, 2010, pp. 31-38.
[136] G. Dubus, "Evaluation of four models for the sonification of elite rowing," *Journal on Multimodal User Interfaces,* vol. 5, pp. 143-156, 2012.
[137] B. Ungerechts, D. Cesarini, and T. Hermann, "Real-time Sonification in Swimming-from pressure changes of displaced water to sound," *BMS2014—Proceedings,* 2014.
[138] T. Hermann, B. Ungerechts, H. Toussaint, and M. Grote, "Sonification of pressure changes in swimming for analysis and optimization," 2012.
[139] M. Ballora, N. A. Giacobe, and D. L. Hall, "Songs of cyberspace: an update on sonifications of network traffic to support situational awareness," in *SPIE Defense, Security, and Sensing*, 2011, pp. 80640P-80640P-6.
[140] D. Ganju and L. Schwiebert, "Using sound for monitoring wireless sensor network behavior," presented at the Proceedings of the 2007 ACM/IFIP/USENIX international conference on Middleware companion, Newport Beach, California, 2007.
[141] M. Ballora, B. Pennycook, P. C. Ivanov, L. Glass, and A. L. Goldberger, "Heart rate sonification: A new approach to medical diagnosis," *Leonardo,* vol. 37, pp. 41-46, 2004.
[142] A. d. Campo, C. Daye, C. Frauenberger, K. Vogt, A. Wallisch, and G. Eckel, "Sonification as an Interdisciplinary Working Process," in *International Conference on Auditory Display*, London, UK, 2006, pp. 28-35.
[143] A. de Campo, "Science By Ear. An Interdisciplinary Approach to Sonifying Scientifc Data," PhD thesis, Institute of Electronic Music and Acoustics-IEM, University for Music and Dramatic Arts Graz, 2009. 4, 21, 22, 2009.
[144] J. Nielsen, "Enhancing the explanatory power of usability heuristics," presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Boston, Massachusetts, USA, 1994.
[145] W. Diaz-Merced, R. Candey, J. Mannone, D. Fields, and E. Rodriguez, "Sonification for the Analysis of Plasma Bubbles at 21 MHz," *Sun and Geosphere,* vol. 3, pp. 42-45, 2008.
[146] D. Merced and W. Liz, "Sound for the exploration of space physics data," University of Glasgow, 2013.
[147] R. M. Candey, A. M. Schertenleib, and W. L. Diaz, "XSonify Sonification Tool For Space Physics," in *International Conference on Auditory Display*, London, UK, 2006, pp. 289-290.
[148] B. N. Walker and J. T. Cothran, "Sonification Sandbox: A Graphical Toolkit for Auditory Graphs," in *International Conference on Auditory Display*, Boston, Ma, 2003, pp. 1-3.
[149] Available: http://www.sonifyer.org/
[150] D. Worrall, M. Bylstra, S. Barrass, and R. Dean, "SoniPy: The design of an extendable software framework for sonification research and auditory display," in *Proceedings of the 13th International Conference on Auditory Display, Montréal, Canada*, 2007.
[151] C. Frauenberger and T. Stockman, "Auditory display design—An investigation of a design pattern approach," *International Journal of Human-Computer Studies,* vol. 67, pp. 907-922, 2009.
[152] J. G. Neuhoff, "Perception, Cognition, and Action in Auditory Displays," in *The Sonification Handbook*, T. Hermann, Neuhoff, J. G. , Hunt, A, Ed., ed Berlin: Logos Verlag, 2011.
[153] Y. L. Bonebright, Flowers, J. H.,, "Evaluation of Auditory Displays," in *The Sonification Handbook*, T. Hermann, Neuhoff, J. G., Hunt, A., Ed., ed, 2011.
[154] D. Brock, J. L. Stroup, and J. A. Ballas, "Effects of 3D Auditory Display on Dual Task Performance in a Simulated Multiscreen Watchstation Environment," *Human Factors and Ergonomics Society Annual Meeting Proceedings,* vol. 46, pp. 1570-1573, 2002.
[155] R. T. H. Laennec, "De l'auscultation médiate ou Traité du diagnostic des maladies des poumons et du coeur, fondé principalement sur ce nouveau moyen d'exploration.," ed, 1816.
[156] E. Du Bois-Reymond, "Über das Telephon," *Archiv für Physiologie,* pp. 573-576, 1877.
[157] A. Volmar, "Listening to the Body Electric. Electrophysiology and the Telephone in the Late 19th Century," *The Virtual Laboratory,* 2010.
[158] L. Hermann, "Ueber electrophysiologische Verwendung des Telephons," *Pflügers Archiv European Journal of Physiology,* vol. 16, pp. 504-509, 1878.

[159] F. Dombois, *The muscle telephone: The undiscovered start of audification in the 1870s*: Max Planck Institute for the History of Science, 2008.
[160] A. d'Arsouval, *Téléphone employé comme galvanoscope*: Gauthier-Villars, 1877.
[161] J. Tarchanow, "Das Telephon als Anzeiger der Nerven-und Muskelströme beim Menschen und den Thieren," *St. Petersburger medicinische Wochenschrift,* vol. 3, pp. 353-54, 1878.
[162] N. E. Wedensky, "Die fundamentalen Eigenschaften des Nerven unter Einwirkung einiger Gifte," *Pflügers Archiv European Journal of Physiology,* vol. 82, pp. 134-191, 1900.
[163] R. Höber, "Ein Verfahren zur Demonstration der Aktionsströme," *Pflügers Archiv European Journal of Physiology,* vol. 177, pp. 305-312, 1919.
[164] F. Scheminzky, "Untersuchungen über die Verstärkung und graphische Registrierung von Schallerscheinungen über Herz und Lunge mittels Elektronenröhren; Konstruktion eines Elektrostethoskops," *Zeitschrift für die gesamte experimentelle Medizin,* vol. 57, pp. 470-501, 1927.
[165] S. Pauletto and A. Hunt, "The sonifiction of EMG data," in *Proceedings of the International Conference Auditory Display (ICAD'06)*, 2006.
[166] S. D. Speeth, *Seismometer Sounds* vol. 33: ASA, 1961.
[167] G. E. Frantti and L. A. Levereault, "Auditory discrimination of seismic signals from earthquakes and explosions," *Bulletin of the Seismological Society of America,* vol. 55, pp. 1-25, February 1, 1965 1965.
[168] G. Frantti and L. Levereault, "Investigation of Auditory Discrimination of Seismic Signals from Earthquakes and Explosions," DTIC Document1964.
[169] K. G. Jansky, "Directional studies of atmospherics at high frequencies," *Proceedings of the Institute of Radio Engineers,* vol. 20, pp. 1920-1932, 1932.
[170] K. G. Jansky, "Electrical disturbances apparently of extraterrestrial origin," *Proceedings of the Institute of Radio Engineers,* vol. 21, pp. 1387-1398, 1933.
[171] A. A. Penzias and R. W. Wilson, "A Measurement of Excess Antenna Temperature at 4080 Mc/s," *The Astrophysical Journal,* vol. 142, pp. 419-421, 1965.
[172] H. Barkhausen, "Two phenomena uncovered with help of the new amplifiers," *Z. Phys,* vol. 20, p. 401, 1919.
[173] H. Barkhausen, "Whistling tones from the Earth," *Proceedings of the Institute of Radio Engineers,* vol. 18, pp. 1155-1159, 1930.
[174] D. A. Gurnett, S. D. Shawhan, N. M. Brice, and R. L. Smith, "Ion Cyclotron Whistlers," *Journal Name: Journal of Geophysical Research (U.S.); Journal Volume: Vol: 70; Other Information: Orig. Receipt Date: 31-DEC-65,* pp. Medium: X; Size: Pages: 1665-88, 1965.
[175] D. Gurnett, R. Shaw, R. Anderson, W. Kurth, and F. Scarf, "Whistlers observed by Voyager 1: Detection of lightning on Jupiter," *Geophysical Research Letters,* vol. 6, pp. 511-514, 1979.
[176] F. Scarf, D. Gurnett, W. Kurth, and R. Poynter, "Voyager plasma wave measurements at Saturn," *Journal of Geophysical Research: Space Physics (1978–2012),* vol. 88, pp. 8971-8984, 1983.
[177] F. Scarf, W. Taylor, C. Russell, and L. Brace, "Lightning on Venus: Orbiter detection of whistler signals," *Journal of Geophysical Research: Space Physics (1978–2012),* vol. 85, pp. 8158-8166, 1980.
[178] W. Kurth, B. D. Strayer, D. A. Gurnett, and F. Scarf, "A summary of whistlers observed by Voyager 1 at Jupiter," *Icarus,* vol. 61, pp. 497-507, 1985.
[179] J. H. Pope, "A high-latitude investigation of the natural very-low-frequency electromagnetic radiation known as chorus," *Journal of Geophysical Research,* vol. 68, pp. 83-99, 1963.
[180] G. Isted and G. Millington, "The 'dawn chorus' in radio observations," *Nature,* vol. 180, 1957.
[181] D. Gurnett, "Electromagnetic plasma wave emissions from the auroral field lines," *Journal of Geomagnetism and Geoelectricity,* vol. 30, pp. 257-272, 1978.
[182] J. M. Watts, "Audio-frequency electromagnetic hiss recorded at Boulder in 1956," *Geofisica pura e applicata,* vol. 37, pp. 169-173, 1957.
[183] R. M. Gallet, "The very low-frequency emissions generated in the earth's exosphere," *Proceedings of the IRE,* vol. 47, pp. 211-231, 1959.
[184] E. J. Smith, R. E. Holzer, and C. T. Russell, "Magnetic emissions in the magnetosheath at frequencies near 100 Hz," *Journal of Geophysical Research,* vol. 74, pp. 3027-3036, 1969.
[185] W. Baumjohann, E. Georgescu, K. H. Fornacon, H. U. Auster, R. A. Treumann, and G. Haerendel, "Magnetospheric lion roars," *Ann. Geophys.,* vol. 18, pp. 406-410, 1999.
[186] C. Higgins, J. Thieman, R. Flagg, F. Reyes, J. Sky, W. Greenman*, et al.*, "Radio Jove: Jupiter Radio Astronomy for Citizens," in *American Astronomical Society Meeting Abstracts*, 2014.
[187] H. Harger and A. Hyde, "Broadcasting The Music Of The Spheres: Creating Radio Astronomy," in *55 th International Astronautical Congress*, 2004.
[188] S. Arnold, "The NASA Radio Jove Project," in *Getting Started in Radio Astronomy*, ed: Springer New York, 2014, pp. 135-167.
[189] J. L. Bougeret, K. Goetz, M. L. Kaiser, S. D. Bale, P. J. Kellogg, M. Maksimovic*, et al.*, "S/WAVES: The Radio and Plasma Wave Investigation on the STEREO Mission," *Space Science Reviews,* vol. 136, pp. 487-528, 2008/04/01 2008.
[190] D. A. Gurnett. (2014). *Space Audio*. Available: http://www-pw.physics.uiowa.edu/space-audio/sounds/

[191] Luxorion. (2013, 12/9/13). *Audio and Sound Files: Geomagnetosphere activity*. Available: http://www.astrosurf.com/luxorion/audiofiles-geomagnetosphere.htm
[192] R. L. Alexander, S. O'Modhrain, J. A. Gilbert, and T. H. Zurbuchen, "Auditory and Visual Evaluation of Fixed-Frequency Events in Time-Varying Signals," in *20th Int. Conf. Aud. Disp (ICAD 2014)*, New York, 2014.
[193] E. Stone, A. Frandsen, R. Mewaldt, E. Christian, D. Margolies, J. Ormes, *et al.*, "The advanced composition explorer," in *The Advanced Composition Explorer Mission*, ed: Springer, 1998, pp. 1-22.
[194] G. B. Andrews, T. H. Zurbuchen, B. H. Mauk, H. Malcom, L. A. Fisk, G. Gloeckler, *et al.*, "The Energetic Particle and Plasma Spectrometer instrument on the MESSENGER spacecraft," *Space Science Reviews,* vol. 131, pp. 523-556, 2007.
[195] G. Gloeckler, J. Cain, F. M. Ipavich, E. O. Tums, P. Bedini, L. A. Fisk, *et al.*, "Investigation of the composition of solar and interstellar matter using solar wind and pickup ion measurements with SWICS and SWIMS on the ACE spacecraft," in *The Advanced Composition Explorer Mission*, ed: Springer, 1998, pp. 497-539.
[196] K. Genuit and R. Sottek, "Application of a New Hearing Model for Determining the Sound Quality of Sound Events.," *Seoul National University, Seoul,* 1995.
[197] D. M. Howard and J. Angus, *Acoustics and psychoacoustics*. Amsterdam ; Boston ; London: Focal Press, 2006.
[198] J. Lee Rodgers and W. A. Nicewander, "Thirteen ways to look at the correlation coefficient," *The American Statistician,* vol. 42, pp. 59-66, 1988.
[199] J. Wagemans, J. H. Elder, M. Kubovy, S. E. Palmer, M. A. Peterson, M. Singh, *et al.*, "A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure–ground organization," 2012.
[200] M. Kubovy and M. Schutz, "Audio-Visual Objects," *Review of Philosophy and Psychology,* vol. 1, pp. 41-61, 2010/03/01 2010.
[201] R. P. Lepping, M. H. Acũna, L. F. Burlaga, W. M. Farrell, J. A. Slavin, K. H. Schatten, *et al.*, "The Wind magnetic field investigation," *Space Science Reviews,* vol. 71, pp. 207-229, 1995.
[202] M. H. Acuña, D. Curtis, J. L. Scheifele, C. T. Russell, P. Schroeder, A. Szabo, *et al.*, "The STEREO/IMPACT Magnetic Field Experiment," in *The STEREO Mission*, C. T. Russell, Ed., ed: Springer New York, 2008, pp. 203-226.
[203] J. Cohen, "Statistical Power Analysis," *Current Directions in Psychological Science,* vol. 1, pp. 98-101, 1992.
[204] J. Cohen, *Statistical power analysis for the behavioral sciences (rev*: Lawrence Erlbaum Associates, Inc, 1977.
[205] R. L. Alexander, S. O'Modhrain, D. A. Roberts, J. A. Gilbert, and T. H. Zurbuchen, "The Bird's Ear View of Space Physics: Audification as a Tool for the Spectral Analysis of Time Series Data," *Journal of Geophysical Research: Space Physics,* p. 2014JA020025, 2014.
[206] K. A. Ericsson and H. A. Simon, *Protocol analysis*: MIT press, 1985.
[207] R. T. Wicks, R. L. Alexander, M. Stevens, L. B. W. III, P. S. Moya, A. Viñas, *et al.*, "A proton cyclotron wave storm generated by unstable proton distribution functions in the solar wind," *The Astrophysical Journal,* vol. (in review), 2015.
[208] R. Alexander. (2015). *Bird's Ear View Dissertation Example Files*. Available: http://deepblue.lib.umich.edu/handle/2027.42/110779
[209] M. E. Fonteyn, B. Kuipers, and S. J. Grobe, "A description of think aloud method and protocol analysis," *Qualitative Health Research,* vol. 3, pp. 430-441, 1993.
[210] C. Lewis, *Using the" thinking-aloud" method in cognitive interface design*: IBM TJ Watson Research Center, 1982.
[211] J. Hughes and S. Parkes, "Trends in the use of verbal protocol analysis in software engineering research," *Behaviour & Information Technology,* vol. 22, pp. 127-140, 2003.
[212] J. Hahn and J. Kim, "Why are some diagrams easier to work with? Effects of diagrammatic representation on the cognitive intergration process of systems analysis and design," *ACM Transactions on Computer-Human Interaction (TOCHI),* vol. 6, pp. 181-213, 1999.
[213] C. B. Seaman, M. G. Mendonça, V. R. Basili, and Y.-M. Kim, "User interface evaluation and empirically-based evolution of a prototype experience management tool," *Software Engineering, IEEE Transactions on,* vol. 29, pp. 838-850, 2003.
[214] T. Hermann, "Sonification for Exploratory Data Analysis," Bielefeld University, 2002.
[215] E. Marsch, R. Schwenn, and B. V. Jackson, "Physics of the inner heliosphere - Part Two - Particles, Waves, and Turbulence," *Solar Physics,* vol. 145, p. 405, 1993.
[216] D. A. Gurnett, *Plasma waves and instabilities* vol. 35: American Geophysical Union, 1985.
[217] L. K. Jian, C. T. Russell, J. G. Luhmann, R. J. Strangeway, J. S. Leisner, and A. B. Galvin, "Ion cyclotron waves in the solar wind observed by STEREO near 1 AU," *The Astrophysical Journal Letters,* vol. 701, p. L105, 2009.

[218] J. K. Edmondson, B. J. Lynch, S. T. Lepri, and T. H. Zurbuchen, "Analysis of High Cadence In Situ Solar Wind Ionic Composition Data Using Wavelet Power Spectra Confidence Levels," *The Astrophysical Journal Supplement Series,* vol. 209, p. 35, 2013.
[219] E. L. Olmsted-Hawala, E. D. Murphy, S. Hawala, and K. T. Ashenfelter, "Think-aloud protocols: a comparison of three think-aloud protocols for use in testing data-dissemination web sites for usability," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2010, pp. 2381-2390.
[220] J. T. Gosling and A. Szabo, "Bifurcated current sheets produced by magnetic reconnection in the solar wind," *Journal of Geophysical Research: Space Physics (1978–2012),* vol. 113, 2008.
[221] R. L. Alexander. (2014). *Birds Ear View: Audification Examples*. Available: http://deepblue.lib.umich.edu/handle/2027.42/106425
[222] E. Landi, R. L. Alexander, J. R. Gruesbeck, J. A. Gilbert, S. T. Lepri, W. B. Manchester, *et al.*, "Carbon Ionization Stages as a Diagnostic of the Solar Wind," *The Astrophysical Journal,* vol. 744, p. 100, 2012.
[223] X. Tang, C. Cattell, L. B. Wilson III, and R. L. Alexander, "First Simultaneous Observations of Lower Hybrid, Whistler-Mode, Electrostatic Solitary, and Electrostatic Electron Cyclotron Waves near the Earth's Magnetopause," presented at the Paper presented at the 2014 American Geophysical Union Conference, San Francisco, 2014.
[224] L. K. Jian, C. T. Russell, R. T. Wicks, M. Stevens, A. Figueroa-Vinas, and R. L. Alexander, "Magnetic waves near the proton cyclotron frequency in the solar wind," in *40th COSPAR Scientific Assembly*, Moscow, Russia, 2014, p. 1331.
[225] R. L. Alexander, J. A. Gilbert, E. Landi, M. Simoni, T. H. Zurbuchen, and D. A. Roberts, "Audification as a Diagnostic Tool for Exploratory Heliospheric Data Analysis," in *17th Int. Conf. Aud. Disp (ICAD 2011)*, Budapest, Hungary, 2011.
[226] *ACE SWICS-SWIMS-V3 Level 2 Data*. Available: http://www.srl.caltech.edu/ACE/ASC/level2/lvl2DATA_SWICS_SWIMS.html
[227] G. Gloeckler, J. Cain, F. M. Ipavich, E. O. Tums, P. Bedini, L. A. Fisk, *et al.*, "Investigation of the composition of solar and interstellar matter using solar wind and pickup ion measurements with SWICS and SWIMS on the ACE spacecraft," *Space Science Reviews,* vol. 86, pp. 497-539, 1998/07/01 1998.
[228] T. H. Zurbuchen, "A new view of the coupling of the Sun and the heliosphere," *Annu. Rev. Astron. Astrophys.,* vol. 45, pp. 297-338, 2007.
[229] J. Geiss, G. Gloeckler, and R. Von Steiger, "Origin of the solar wind from composition data," *Space Science Reviews,* vol. 72, pp. 49-60, 1995.
[230] Y.-K. Ko, L. A. Fisk, J. Geiss, G. Gloeckler, and M. Guhathakurta, "An empirical study of the electron temperature and heavy ion velocities in the south polar coronal hole," *Solar Physics,* vol. 171, pp. 345-361, 1997.
[231] R. Von Steiger, J. Geiss, and G. Gloeckler, "Composition of the solar wind," in *Cosmic winds and the heliosphere*, 1997, p. 581.
[232] Jim Raines, S. Lepri, and T. Zurbuchen. (2013, December 04). *SWICS/SWIMS Level 2 Release Notes, Data Version 3.3.1*. Available: http://www.srl.caltech.edu/ACE/ASC/DATA/level2/ssv3/swics_lv2_v3_release_notes.txt
[233] H. Newton and M. Nunn, "The sun's rotation derived from sunspots 1934–1944 and additional results," *Monthly Notices of the Royal Astronomical Society,* vol. 111, pp. 413-421, 1951.
[234] J. Lei, J. P. Thayer, J. M. Forbes, E. K. Sutton, and R. S. Nerem, "Rotating solar coronal holes and periodic modulation of the upper atmosphere," *Geophysical Research Letters,* vol. 35, p. L10109, 2008.
[235] R. Schwenn and E. Marsch, *Physics of the Inner Heliosphere. I. Large Scale Phenomenon* vol. Vol. 1. Heidelberg, Germany: Springer-Verlag, 1990.
[236] H. B. Snodgrass and R. K. Ulrich, "Rotation of Doppler features in the solar photosphere," *The Astrophysical Journal,* vol. 351, pp. 309-316, 1990.
[237] T. H. Zurbuchen, S. Hefti, L. A. Fisk, G. Gloeckler, and N. A. Schwadron, "Magnetic structure of the slow solar wind: Constraints from composition data," *Journal of Geophysical Research: Space Physics,* vol. 105, pp. 18327-18336, 2000.
[238] M. Tervaniemi, V. Just, S. Koelsch, A. Widmann, and E. Schröger, "Pitch discrimination accuracy in musicians vs nonmusicians: an event-related potential and behavioral study," *Experimental brain research,* vol. 161, pp. 1-10, 2005.
[239] R. L. Alexander, T. H. Zurbuchen, J. A. Gilbert, S. T. Lepri, and J. M. Raines, "Sonification of ACE Level 2 Solar Wind Data," presented at the 16th Int. Conf. Aud. Disp (ICAD 2010), Washington DC, 2010.
[240] L. B. Wilson III, "The microphysics of collisionless shocks," University of Minnesota, 2010.
[241] J. W. Bonnell, F. S. Mozer, G. T. Delory, A. J. Hull, R. E. Ergun, C. M. Cully, *et al.*, "The Electric Field Instrument (EFI) for THEMIS," *Space Science Reviews,* vol. 141, pp. 303-341, 2008/12/01 2008.
[242] V. Angelopoulos, "The THEMIS Mission," in *The THEMIS Mission*, J. L. Burch and V. Angelopoulos, Eds., ed: Springer New York, 2009.
[243] C. M. Cully, R. E. Ergun, K. Stevens, A. Nammari, and J. Westfall, "The THEMIS Digital Fields Board," *Space Science Reviews,* vol. 141, pp. 343-355, 2008/12/01 2008.

[244] L. Wilson, D. Sibeck, A. Breneman, O. L. Contel, C. Cully, D. Turner, *et al.*, "Quantified energy dissipation rates in the terrestrial bow shock: 1. Analysis techniques and methodology," *Journal of Geophysical Research: Space Physics,* vol. 119, pp. 6455-6474, 2014.
[245] L. Wilson, D. Sibeck, A. Breneman, O. L. Contel, C. Cully, D. Turner, *et al.*, "Quantified energy dissipation rates in the terrestrial bow shock: 2. Waves and dissipation," *Journal of Geophysical Research: Space Physics,* vol. 119, pp. 6475-6495, 2014.
[246] H. Matsumoto, H. Kojima, T. Miyatake, Y. Omura, M. Okada, I. Nagano, *et al.*, "Electrostatic solitary waves (ESW) in the magnetotail: BEN wave forms observed by GEOTAIL," *Geophysical Research Letters,* vol. 21, pp. 2915-2918, 1994.
[247] N. Singh, S. M. Loo, and B. E. Wells, "Electron hole as an antenna radiating plasma waves," *Geophysical Research Letters,* vol. 28, pp. 1371-1374, 2001.
[248] N. Singh, S. M. Loo, B. E. Wells, and C. Deverapalli, "Three-dimensional structure of electron holes driven by an electron beam," *Geophysical Research Letters,* vol. 27, pp. 2469-2472, 2000.
[249] L. P. Dyrud and M. M. Oppenheim, "Electron holes, ion waves, and anomalous resistivity in space plasmas," *Journal of Geophysical Research: Space Physics,* vol. 111, p. A01302, 2006.
[250] M. V. Goldman, D. L. Newman, G. Lapenta, L. Andersson, J. T. Gosling, S. Eriksson, *et al.*, "Emission of Quasi parallel Whistlers by Fast Electron Phase-Space Holes during Magnetic Reconnection," *Physical Review Letters,* vol. 112, p. 145002, 04/08/ 2014.
[251] J. G. Proakis, *Digital signal processing: principles algorithms and applications*: Pearson Education India, 2001.
[252] J. D. Richardson and C. W. Smith, "The radial temperature profile of the solar wind," *Geophysical research letters,* vol. 30, 2003.
[253] P. A. Isenberg, "A self-consistent marginally stable state for parallel ion cyclotron waves," *Physics of Plasmas (1994-present),* vol. 19, p. 032116, 2012.
[254] M. L. Goldstein, D. A. Roberts, and W. Matthaeus, "Magnetohydrodynamic turbulence in the solar wind," *Annual review of astronomy and astrophysics,* vol. 33, pp. 283-326, 1995.
[255] S. P. Gary, *Theory of space plasma microinstabilities*: Cambridge university press, 2005.
[256] S. P. Gary, C. W. Smith, M. A. Lee, M. L. Goldstein, and D. W. Forslund, "Electromagnetic ion beam instabilities," *Physics of Fluids (1958-1988),* vol. 27, pp. 1852-1862, 1984.
[257] J. C. Kasper, A. J. Lazarus, and S. P. Gary, "Wind/SWE observations of firehose constraint on solar wind proton temperature anisotropy," *Geophysical research letters,* vol. 29, pp. 20-1-20-4, 2002.
[258] P. Hellinger, P. Trávníček, J. C. Kasper, and A. J. Lazarus, "Solar wind proton temperature anisotropy: Linear theory and WIND/SWE observations," *Geophysical research letters,* vol. 33, 2006.
[259] L. Matteini, S. Landi, P. Hellinger, F. Pantellini, M. Maksimovic, M. Velli, *et al.*, "Evolution of the solar wind proton temperature anisotropy from 0.3 to 2.5 AU," *Geophysical Research Letters,* vol. 34, 2007.
[260] S. Bale, J. Kasper, G. Howes, E. Quataert, C. Salem, and D. Sundkvist, "Magnetic fluctuation power near proton temperature anisotropy instability thresholds in the solar wind," *Physical review letters,* vol. 103, p. 211101, 2009.
[261] S. Servidio, K. Osman, F. Valentini, D. Perrone, F. Califano, S. Chapman, *et al.*, "Proton kinetic effects in Vlasov and solar wind turbulence," *The Astrophysical Journal Letters,* vol. 781, p. L27, 2014.
[262] R. T. Wicks, L. Matteini, T. S. Horbury, P. Hellinger, and D. A. Roberts, "Temperature anisotropy instabilities; combining plasma and magnetic field data at different distances from the Sun," *AIP Conference Proceedings,* vol. 1539, pp. 303-306, 2013.
[263] L. Jian, H. Wei, C. Russell, J. Luhmann, B. Klecker, N. Omidi, *et al.*, "Electromagnetic Waves near the Proton Cyclotron Frequency: STEREO Observations," *The Astrophysical Journal,* vol. 786, p. 123, 2014.
[264] L. Jian, C. Russell, J. Luhmann, B. Anderson, S. Boardsen, R. Strangeway, *et al.*, "Observations of ion cyclotron waves in the solar wind near 0.3 AU," *Journal of Geophysical Research: Space Physics (1978–2012),* vol. 115, 2010.
[265] K. Ogilvie, D. Chornay, R. Fritzenreiter, F. Hunsaker, J. Keller, J. Lobell, *et al.*, "SWE, a comprehensive plasma instrument for the Wind spacecraft," *Space Science Reviews,* vol. 71, pp. 55-77, 1995.
[266] M. Wiedenbeck, G. Mason, R. Gómez-Herrero, D. Haggerty, N. Nitta, C. Cohen, *et al.*, "Observations of a 3He-rich SEP Event over a Broad Range of Heliographic Longitudes: Results from STEREO and ACE," in *Twelfth International Solar Wind Conference*, 2010, pp. 621-624.
[267] *CDAWeb*. Available: http://cdaweb.gsfc.nasa.gov/cdaweb/sp_phys/
[268] H. Nyquist, "Certain topics in telegraph transmission theory," *American Institute of Electrical Engineers, Transactions of the,* vol. 47, pp. 617-644, 1928.
[269] A. Pras and C. Guastavino, "Sampling Rate Discrimination: 44.1 kHz vs. 88.2 kHz," in *Audio Engineering Society Convention 128*, 2010.
[270] E. W. Start, "Direct sound enhancement by wave field synthesis," 1997.
[271] M. Pivec, *Affective and emotional aspects of human-computer interaction: game-based and innovative learning approaches* vol. 1: IOS Press, 2006.

[272] R. E. Clark and B. M. Sugrue, "Research on instructional media, 1978-1988," *Educational media and technology yearbook,* vol. 14, pp. 19-36, 1988.
[273] K. H. Brodersen, C. S. Ong, K. E. Stephan, and J. M. Buhmann, "The balanced accuracy and its posterior distribution," in *Pattern Recognition (ICPR), 2010 20th International Conference on*, 2010, pp. 3121-3124.
[274] R. W. Peeling, P. G. Smith, and P. M. M. Bossuyt, "A guide for diagnostic evaluations," *Nat Rev Micro*.
[275] R. L. Alexander, O'Modhrain, S., Gilbert, J.A. Zurbuchen, T.H., Simoni, M., "Recognition of Audified Data in Untrained Listeners," presented at the International Conference on Auditory Display, Georgia Institute of Technology, 2012.
[276] C. J. Lintott, K. Schawinski, A. Slosar, K. Land, S. Bamford, D. Thomas*, et al.*, "Galaxy Zoo: morphologies derived from visual inspection of galaxies from the Sloan Digital Sky Survey," *Monthly Notices of the Royal Astronomical Society,* vol. 389, pp. 1179-1189, September 21, 2008 2008.
[277] A. Balogh, T. J. Beek, R. Forsyth, P. Hedgecock, R. Marquedant, E. Smith*, et al.*, "The magnetic field investigation on the ULYSSES mission-Instrumentation and preliminary scientific results," *Astronomy and Astrophysics Supplement Series,* vol. 92, pp. 221-236, 1992.
[278] A. Supper, "Sublime frequencies: The construction of sublime listening experiences in the sonification of scientific data," *Social Studies of Science,* vol. 44, pp. 34-58, 2014.
[279] S. Dow, B. MacIntyre, J. Lee, C. Oezbek, J. D. Bolter, and M. Gandy, "Wizard of Oz support throughout an iterative design process," *Pervasive Computing, IEEE,* vol. 4, pp. 18-26, 2005.