LISTENER, TASK, AND AUDITORY GRAPH: TOWARD A CONCEPTUAL MODEL OF AUDITORY GRAPH COMPREHENSION

Michael A. Nees & Bruce N. Walker

Sonification Lab, School of Psychology
Georgia Institute of Technology
Atlanta, GA, USA 30332
mnees@gatech.edu; bruce.walker@psych.gatech.edu

ABSTRACT

Auditory graph design and implementation often has been subject to criticisms of arbitrary or atheoretical decision-making processes in both research and application. Despite increasing interest in auditory displays coupled with more than two decades of auditory graph research, no theoretical models of how a listener processes an auditory graph have been proposed. The current paper seeks to present a conceptual level account of the factors relevant to the comprehension of auditory graphs by human listeners. We attempt to make links to the relevant literature on basic auditory perception, and we offer explicit justification for, or discussion of, a number of common design practices that are often justified only implicitly or by intuition in the auditory graph literature. Finally, we take initial steps toward a qualitative, conceptual level model of auditory graph comprehension that will help to organize the available data on auditory graph comprehension and make predictions for future research and applications with auditory graphs

[Keywords: Auditory Graphs, Sonification, Psychological Models, Individual Differences, Task]

1. INTRODUCTION

Sonification, the use of nonspeech audio as a means of information display [1], is a multidisciplinary approach to auditory information display that lies at the intersection of such diverse domains as psychology, audiology, music, and computer science. As a result of the multifaceted nature of sonification, the work of the field does not fit neatly into the existing theoretical frameworks of any of its constituent domains. Recently it has been suggested that many decisions regarding the design and application of sonifications are made arbitrarily [2, 3] and that much of the knowledge generated from sonification research may be difficult to generalize beyond the narrow specifications contrived in any particular study of sonification [4].

Indeed, the field of sonification has been slow to develop a cohesive account of how people interact with auditory representations of information, and this lack of an organizational framework has resulted in some arbitrary design decisions that have not followed a formal logic. While a wealth of valuable sonification research has been conducted and published in recent years, even a generic framework for the organization of this information has yet to be established [but for a recent development, see 5]. Furthermore, the links and mutually beneficial relationships between the sonification literature and existing theory and literature in other areas of science, which certainly exist, have not been convincingly established. The present sonification literature, however, could offer insights and contributions for theory from related fields, just as auditory display research has benefited (both implicitly and explicitly) from approaches to information display from other disciplines.

As an example, sonification research has undoubtedly benefited from the decades of work on basic auditory perception, but this heritage is often implicit in the sonification literature. Design decisions in the sonification literature are often not explicitly justified, and this likely fuels criticism that the field operates on ad hoc, arbitrary principles when building sound displays. Even in the absence of a stated rationale, design decisions may be reasonable and informed with regards to the basic properties of auditory perception, and some consistencies have emerged across different studies and applications.

The current paper attempts to make a few links between sonification research and relevant theoretical approaches in other fields, especially psychology. While a comprehensive review of all related topics is beyond the scope of this paper, we argue that sonification research has not and does not take place in a theoretical vacuum, but instead has borrowed from, and can contribute back to, existing theoretical approaches in a number of other fields. We contend that sonification researchers already possess the basic building blocks of knowledge to begin to articulate high-level accounts of how humans interact with auditory displays. Our scope is constrained to a conceptual level model of how people comprehend the information presented in auditory graphs—a class of sonifications that use sounds to represent quantitative data—but our approach may generalize to other classes of sonifications. We believe that this approach represents the initial steps toward formally describing a theory for auditory graphs that organizes the knowledge derived from existing research and makes some broad and basic predictions for future research and applications of auditory graphs.

2. RATIONALE AND SCOPE

Graphical representations of information are pervasive in the publications of both science and popular culture [see, for example, 6]. Visual graphs can offer a concise summary of data (relative to other presentation formats such as text, tables, etc.). Graphs may also offer the user effortless access to some data features (e.g., patterns) that are not immediately evident in non-graphical representations of the same data [see 7].

The pervasiveness and utility of graphs have lead researchers to examine the potential to present graphical information with sound. Sound has obvious potential as an assistive technology for visually impaired people, but there are numerous tasks and environments where sound displays may also benefit sighted listeners. For example, research has shown that auditory displays can be useful in scenarios where the display is small, such as with mobile computing devices [see 8, 9], or when the visual system may be overtaxed by traditional displays [10, 11].

Auditory graphs, then, are a class of sonified displays that use sound to represent quantitative information. Within the framework of de Campo's [5] sonification design space map, auditory graphs employ parameter mapping sonification techniques. In other words, changes in quantitative data are mapped to changes in a dimension or multiple dimensions of sound [see 12]. Auditory graphs have most commonly mapped changes in quantitative data to changes in the frequencies of sounds in time, thus frequency and time have been equated with traditional Y- and X-axes in visual graphs. Although the reasons for using frequency mappings in auditory graphs are often not made explicit or are justified only intuitively, a wealth of theory on pitch perception may support the intuition that frequency offers the best choice of a sound dimension mapping for quantitative data in auditory graphs. Shepard's [see, for example, 13] helix representation of pitch perception described the perception of frequency in spatial terms, with higher frequencies represented as higher in space moving up the helix. Kubovy's Theory of Indispensable Attributes [14] offered another possible theoretical explanation for the choice of frequency changes as a primary mapping for data changes in auditory graphs, as he proposed that "visual spatial location is analogous to auditory frequency" (pp. 78). While it is unlikely that any single theory or researcher inspired the common practice of frequency mapping in auditory graphs, the convergence of insights from multiple theories of pitch perception suggests that frequency is an appropriate auditory analog for visual Y-axis space in Cartesian coordinates. Not surprisingly, then, auditory graph researchers have repeatedly recommended that quantitative (i.e., graphical Y-axis) values be mapped to frequency [15-17], and this practice pervades auditory graph literature.

We limit our discussion here to those sonifications that use frequency mapping in time as a primary means of data display. We also limit the scope of our discussion to those data sets whose size and dimensionality make parameter mapping the most appropriate sonification technique [see 5]. Our concern here is with auditory representations of data that are appropriate for traditional graphical displays such as the line graphs, scatterplots, histograms, etc., which pervade popular media and scientific publications. Researchers have suggested [e.g. 18, 19], and we contend here, that auditory representations of even simple graphical displays (i.e., the bivariate and multivariate plots found in popular media and scientific journals) can be helpful for the educational and data exploration needs of both sighted and visually impaired persons.

3. TOWARD A MODEL OF AUDITORY GRAPH COMPREHENSION

Visual graph comprehension literature is replete with models and theories that attempt to explain the respective roles of perceptual and cognitive factors, individual differences, and task dependencies in the comprehension of graphs [e.g., 7, 20, 21-23], but no similar organizing framework has attempted to account for the importance of similar or equivalent variables in auditory graph comprehension. While theories of visual graph comprehension will not necessarily translate directly to the auditory domain, they can offer a useful starting point in the formulation of a model of auditory graph comprehension.

The proposed model of auditory graph comprehension is intended to assimilate relevant insights from visual graph comprehension literature with knowledge from other relevant areas of study (e.g., auditory graphs, auditory perception, and music) to arrive at a plausible explanation of the factors that influence the comprehension of auditory graphs.

Of note, *graph comprehension* refers to the extent to which a human listener is able to extract the information desired (as defined by her or his task) from the display. The construct of comprehension is often operationally defined and measured as relevant dependent variables in auditory graph research, most often accuracy and response time. We emphasize, then, the contributions of independent variables related to the task, the listener, and the auditory graph display to auditory graph comprehension. We first discuss each of these groups of variables separately, but in section 3.4 we emphasize the interactive and mutually influential nature of combinations of these variables on auditory graph comprehension.

3.1. Task

The task dependent nature of human interactions with displays has been widely discussed with regards to both visual graphs [see 23, 24] and auditory displays and graphs [see, for example, 25, 26, 27]. Given that the function of sonification is to convey information [1], we propose that, for auditory graphs, the listener's *task* defines and constrains the information the human listener wishes to extract from the auditory graph. In other words, the task predetermines the information the listener seeks.

Tasks with graphs can be as broad as general data exploration or as narrow as point estimation. Formal task analysis methods have been described to break down a given task into its component parts [see, for example, 28], and these methods have been applied to some extent in the realm of auditory graphs [e.g., 29]. Barrass [30] discussed information-seeking with auditory displays at great length and offered an extended discussion of task questions and purposes. A task analytic approach may help to categorize the processes involved in extracting information from the display.

Cleveland and McGill [31] proposed a theory that ranked "elementary perceptual tasks" (p. 531) of visual graph information-seeking in order of difficulty. Although no comparable categorization of tasks has been proposed for auditory graphs, Jones' [32] rhythmic theory of auditory pattern perception offered insights regarding the type of information that can be communicated in a sequence of sounds. She suggested that sounds presented in time could be described as amenable to the perception of nominal, ordinal, and interval relationships. With nominal relationships between sounds, the listener may only be able to perceive that the sounds of a sequence are the same or different. Ordinal relations allow for the perception of direction, with one sound having higher or lower frequency than its comparator. Finally, interval relations allow for the perception of both direction and magnitude of frequency differences. We propose that Jones' rhythmic theory represents an important starting point in predicting the difficulty of auditory graphing tasks. Tasks requiring information regarding only nominal relations should be easier for listeners to perform than tasks that require ordinal relations, which in turn should be easier than tasks that require information regarding interval relations. Research has suggested that trend tasks with auditory graphs—which generally require ordinal information regarding the direction of pitch changes—are readily accomplished [33]. Point estimation tasks, which require judgments regarding both direction and magnitude of pitch change, have proven to be particularly difficult in studies to date [see 29, 34, 35, 36], which is in accordance with Jones' predictions.

3.2. Listener Characteristics

Research in basic auditory perception has established a foundation of knowledge regarding the lawfulness of auditory sensation and perception, yet each human listener of an auditory graph possesses a unique set of capabilities, limitations, and proclivities that may influence her or his ability to successfully accomplish a task with a given display.

3.2.1. Commonalities

A number of properties of the biological apparatus for hearing can be described more or less lawfully across individuals [e.g., thresholds, etc., see 37], particularly in the range of frequencies and amplitudes employed in auditory graph design.

For example, detection of the sound is a necessary prerequisite for the comprehension of an auditory display, and the fundamental capabilities of the auditory system are fairly well understood and establish limits for the detectability of sounds along a number of dimensions, including frequency and amplitude. Although upper and lower limits of detectability along a given sound dimension are to be avoided in auditory display design, auditory display researchers have made the logical step beyond simple detectability and chosen design parameters that exploit knowledge regarding the most sensitive response regions of the auditory apparatus along the dimensions of sound to which data are mapped. Although a complete discussion of the general capabilities and limitations of auditory sensation are beyond the scope of this paper [for a more complete discussion, see, for example, 37, 38], auditory graph designers have proceeded with an awareness of the basic lawfulness of auditory sensation.

Equally as important to the current discussion are the lawful properties of auditory perception, whereby sounds are organized into a meaningful representation of the world. Kosslyn theorized that in visual graphs, some patterns may be perceived automatically. Others have described this general phenomena in perception as a function of emergent features or preattentive processing [see 39, 40], default encoding [20], and Gestalts [41, 42], but the underlying concept in each instance emphasizes relatively effortless and error-free information extraction. Bregman [43] has similarly described a process whereby schemas—well-learned patterns—can be automatically activated to aid in the recognition of auditory patterns. Such data features or emergent patterns in auditory graphs, when present, should be extracted automatically and easily comprehended.

Auditory pattern perception theories have made concrete predictions regarding the exact properties of auditory patterns that should make them amenable to easier information extraction. Jones [32] predicted how relationships between tones specify the complexity of auditory patterns. Simple patterns—those patterns that are most amenable to processing interval relationships (i.e., both the direction and magnitude of frequency changes)—were theorized to be those that had regular, monotonic changes in frequency with regular timing (i.e., frequencies increasing or decreasing at regular pitch and time intervals). Sequences of tones with contour changes and irregular interval changes (in pitch or time) were more complex in the theory, and perhaps only amenable to perceptions of ordinal relations in some circumstances. Finally, in cases with very large interval jumps (i.e., large changes in frequency from tone to tone) and multiple changes in direction, the perception of a temporal sequence may collapse altogether. In this instance, the listener may report two streams of tones grouped as high and low in frequency rather than a perception of a single stream of alternating tones [also see 44], and in extreme cases only nominal relations between the sounds will be perceived. Jones posited that auditory patterns are represented as nested hierarchies of tone sequences, with more complex sequences breaking down into smaller parts.

Deutsch and Feroe [45] proposed a similar hierarchical model of the perception of musical tone sequences. They suggested that the initial perceptual grouping of tone sequences is determined by Gestalt principles like proximity (in either frequency or time) and good continuation (such as when a series of pitches consistently increase), and the model predicts that patterns exhibiting better Gestalt groupings are easier to perceive. Deutsch and Feroe also suggested that large jumps in frequency intervals from tone to tone are detrimental to grouping. Vickers and Hogg [46] recently proposed a design space that emphasized the similarities between sonifications and musical compositions, and they suggested that the most successful sonifications will be those whose properties allow the listener to "attend carefully" to the data rendering. Theories of auditory pattern perception [e.g., 32] and musical sequence perception [45] may offer predictions regarding the patterns of tones in time that allow for the most easily extractable Gestalts or emergent features to be perceived.

3.2.2. Individual Differences

Despite the utility of knowledge regarding the lawfulness of auditory sensation and perception, auditory graph comprehension undoubtedly calls upon other perceptual and cognitive capabilities and limitations that exhibit relatively high variability from person to person. Although the importance of individual differences in graph comprehension (both auditory and visual) is not well understood, we posit here a number of individual difference variables that should be important predictors of auditory graph comprehension.

Musical ability, often measured as a function of musical experience or training, has been frequently posited as an important contributor to performance with auditory displays. Indeed, in some instances researchers have found differences in performance with sonifications that favor musicians over nonmusicians [e.g., 34, 47]. Other research in the field has found either no differences based on this variable or no predictive power of musical experience [for a brief review, see 48]. Perhaps part of the discrepancy lies in the lack of a valid metric for assessing musical ability per se, as the use of surrogate measures (e.g., years of musical training) may not fully capture individual differences in musical *ability*. Some work has suggested that musical ability likely does not play an important predictive role in performance with auditory displays, as both musicians and nonmusicians may exhibit keen auditory perceptual abilities [see 48, 49]. The currently available knowledge suggests that musical experience (as a surrogate for musical ability) offers little predictive power for performance with auditory displays.

Recently, research has begun to examine the role of cognitive abilities in the comprehension of graphs. Trickett and Trafton [51] have argued that graph comprehension literature has overlooked the role of spatial abilities. They suggest that spatial cognition plays an important role in understanding graphs, particularly when the information needed is not immediately available from perceptual processes. Spatial abilities likely will also play a role in auditory graph comprehension, as both theory [e.g., 13] and empirical evidence [e.g., 52] have suggested that pitch can be represented spatially. Toth and Lewis [53] found that in some circumstances verbal memory impacted graph comprehension. Although verbal memory has not been investigated with regards to auditory graphs, it may play an important role given that much of the contextual information (data labels, etc.) in auditory graphs typically has been presented in text accompanying the sounds. Individual differences in domain knowledge have also been posited to play a role in auditory graph comprehension [see 54]. Though little work has attempted to understand how individual differences impact auditory display performance, Walker and Mauney [50] found evidence that Raven's Progressive Matrices helped to predict performance on a magnitude estimation sonification task. Clearly, however, much more work is needed to clarify the cognitive variables relevant to auditory graph comprehension.

Walker and Lane [55] found that preferred polarities for mapping sound to conceptual data dimension, discussed in more detail in the next section, were sometimes reversed for visually impaired as compared to sighted listeners. Their data generally indicated similarities and consensus in perception across those populations. The few exceptions where visually impaired and sighted users reported opposite polarities, suggest that auditory graph designers should be aware of the potential for differences along this important population variable, especially if the intended audience of the auditory graph includes visually impaired listeners.

3.2.3. Training and Learning

Auditory graph studies to date have generally sampled from populations of naïve listeners who have never heard data before entering a laboratory to participate in a study. As a result, the ceiling of performance for trained auditory graph listeners remains entirely unknown. Visual graph viewers typically have years of formal training and informal experience with visual information displays that employ Cartesian coordinates. Until longitudinal studies of learning are conducted, the full potential of auditory graphs will be unclear both in absolute terms and relative to the efficacy of traditional visual graphs. Smith and Walker [29] found, not surprisingly, that a brief training period improved immediate performance of a point estimation task with auditory graphs. Walker and Nees [35] found that either a brief instructional training program or practice with feedback regarding correct responses improved performance of the point estimation task, but simple repeated exposure to the task (in the absence of feedback) did not. Clearly, these studies represent only an initial starting point for describing the course of skill acquisition with auditory graphs. Future research should disentangle the effects of explicit training versus simple practice with or informal exposure to auditory graphs. Longitudinal studies should examine the time course of learning in auditory graphs and offer valuable data regarding skilled users' performance with the displays. We suggest here that the data currently available for performance with auditory graphs may very well underestimate the potential usefulness of auditory graphs for listeners who develop extensive skill and experience with the displays.

3.3. Auditory Graph Display Characteristics

The final major grouping of variables in our discussion has been reserved for the design of the auditory graph, which represents the actual bottom-up stimulus for our model of auditory graph comprehension. The display represents perhaps the most researched part of our model, as many different techniques and approaches to building auditory graphs have been attempted, and guidelines for designing auditory graphs have been offered [16].

3.3.1. Data

The auditory graph display design begins with some quantitative data that are to be represented. The nature of the data and its qualities may play an important role in auditory graph comprehension. Many data sets are finite and known before a rendering in sound is produced, in which case information regarding known maxima, minima, means, etc., may be incorporated into the design of the display. Other data, however, may be of interest as they occur in real time [see 56], which may present unique challenges for the auditory graph designer. Barrass [30] called for display designers to pay heed to the nature of the data to be represented when designing auditory displays, and he offered a method for the formal examination and characterization of data during the display design process.

3.3.2. Mappings, Scalings, and Polarities

Decisions regarding mappings, scalings, and polarities are especially critical in the design of auditory graphs. Walker [12, 57] has examined these issues in detail. Mapping refers to the dimension of sound that is chosen to covary with changes in the data represented in auditory graphs. As mentioned above, changes in data are most often mapped to changes in the frequency of sounds in auditory graphs. Although other mappings for sonification have been examined [e.g., tempo, brightness, etc., see 12, 57] and carefully chosen redundant or dual mappings may be desirable[1] [25], we limit our discussion here to frequency mappings. As was mentioned above, frequency mappings are common in auditory graphs and relatively robust, and we can apply existing theoretical frameworks [e.g., 32, 45] to make concrete predictions regarding the perceivability of patterns of tones that change in frequency over time. For the frequency mapping used in auditory graphs, a display designer must also choose the type of sound (pure tones, MIDI instruments, etc.) that will be mapped to the data. Brown et al. [16] argued that MIDI instruments should be employed, as past research found that musical instruments were more pleasant and easier to perceive than pure tones [59]. Auditory graph researchers have often used the MIDI piano timbre as a primary or first option for mapping, probably owing to the large range of frequencies naturally spanned by the piano [see 38].

Scaling refers to the amount of change in a sound dimension used to represent a unit of change in the conceptual data dimension being represented. Walker [12, 57] used magnitude estimation to determine the preferred scaling slopes for mapping

---

[1] Some redundant mappings may be detrimental to performance with auditory graphs. One study [58] attempted to use panning and spatial elevation as well as frequency to represent data in auditory graphs, which resulted in exceptionally poor performance in the reproduction of simple linear increasing functions.

frequency to a number of conceptual data dimensions; interestingly, the scaling slope for the same quantitative changes in data may be different depending upon the conceptual data dimension (e.g., temperature, size, pressure, etc.) being represented. When possible, the scaling factor for an auditory graph should match the preferred scaling for the conceptual dimension represented. Brown et al. [16] recommend that scaling of data in auditory graphs should not exceed a minimum of MIDI note 35 (~61.7 Hz) or a maximum of MIDI note 100 (~2637 Hz). They cite difficulties in hardware (i.e., computer sound card) reproduction of notes outside of this range, but it is also important to note that the natural range of many musical instruments fall within or are centered within this range [see 38].

Another important, basic consideration for auditory graph design is the polarity of the mapping between frequency and the data represented. Convention has shown that *higher* frequencies should represent *more* of a conceptual data dimension, while lower frequencies should correspond to lower quantities of the data dimension [16], and research [12, 57] has shown that this positive polarity mapping (increasing-increasing) is intuitive for most listeners across most data dimensions. As mentioned above, however, Walker and Lane [55] found that polarities for visually impaired as compared to sighted listeners were sometimes reversed, whereby increasing frequencies intuitively mapped to less of a particular data dimension. While positive polarities for frequency mappings should generally match intuitive listener preferences, designers should be careful to note and take account of exceptions to this rule. Auditory graphs whose polarities oppose intuitive mappings with regards to the direction of increase or decrease may be harder to comprehend.

3.3.3. *Context*

Graphical context describes those aspects of the display beyond the actual data representation that are included to facilitate comprehension of data relationships. Seminal graph comprehension theory was primarily concerned with those stimulus dimensions used to represent the data [e.g., 31], but later graph theory was expanded to include other essential parts of the graph such as the background, axes, and labels [7]. Empirical research has confirmed the importance of contextual information to the understanding of a graph [22]. Much like early theory and research in visual graph comprehension, auditory graph research generally has been concerned with the actual dimensions of sound used to specify data (e.g., simple tone sequences). Recently, however, researchers have begun to consider ways to establish context and frame the data in an auditory graph.

Smith and Walker [29, 60] have shown that Y-axis context in the form of reference tones can facilitate performance of a point estimation task with auditory graphs. Likewise, Smith and Walker found that X-axis context in the form of rhythmic clicks or beats [also see 61] was generally helpful. Of note, as more concurrent sounds are added to the display, issues of masking become important. Context has generally been implemented using timbres that are distinct from the data timbre (typically a piano instrument). Nees and Walker [36] offered evidence that relative intensity adjustments may also facilitate the perceptual segregation of data from context in auditory graphs.

These findings suggest that concurrent auditory context may be an important aid to some tasks with auditory graphs. The role of context, however, is not well understood, especially with regard to the interactions of context with user variables (e.g., training) and task dependencies. Furthermore, context can sometimes be provided to auditory graph listeners without introducing concurrent sounds to the display. A taxonomy of graphical context would include declarative context, such as verbal or textual instructions regarding the mapping and scaling of the data, etc., as well as scaling cues like reference tones that can be presented before (as opposed to concurrent with) the actual data tones. More research is needed to clarify the appropriate use of context in auditory graphs, but contextual cues should play an important role in auditory graph comprehension for a number of listeners and tasks.

3.3.4. *Temporal characteristics of auditory graph stimuli*

The issues surrounding the temporal characteristics of auditory graphs are numerous and probably intricately inter-related with each other and with other variables discussed throughout this paper. Research has yet to determine the ideal or appropriate permissible lengths (in time) of auditory graph stimuli. Flowers et al. [26] suggested that, for some tasks, auditory graphs of durations under 10 seconds might be appropriate given the length of time that auditory sensory information can be stored. They point out, however, that data and task characteristics will also play a role in determining the ideal time frame for presentations of auditory graphs.

Another important temporal consideration in auditory graph design involves the ideal rate of presentation or data density (i.e., the number of tones presented per second). Nees [33] recently found little effect of presentation rates (ranging from 1 data point per second up to 8 data points per second) on a trend identification task, but there was a small effect of presentation rate on a point estimation task whereby performance was better with either 1 or 4 data points per second. This topic, however, requires more research to determine the unique contribution of presentation rate to auditory graph comprehension.

Of note, increasing the presentation rate necessarily decreases the amount of time that can be occupied by individual, discrete tones. Early work suggested that pitch perception deteriorated as the duration of individual tones fell below 100 ms [62]; later research showed that this effect was dependent upon both the frequency and intensity of the tone and possibly only of great consequence at much shorter tonal durations [e.g., around 25 ms, see 63]. Similarly, research on the perception of numerosity with tones (i.e., accurately perceiving how many tones are heard) has shown that accuracy decreases as both the rate of presentation and the overall number of tones increase [64], with a possible perceptual ceiling at around 9 to 11 sounds per second [65]. Depending upon the task, the fallibility of both pitch perception at shorter tonal durations and the perception of numerosity at high presentation rates may set an upper limit for the number of discrete data points that can be presented per second in auditory graphs. Brown et al. [16] recommended allowing at least 50-70 ms between tones in auditory graphs in order to ensure the data are comprehensible.

3.3.5. *Multiple Data Series*

Thus far our discussion, at least implicitly, has concerned only auditory graphs that present a single data series. Auditory graph researchers have experimented with methods for presenting multiple data series within the same auditory plot. Bonebright et al. [61] used a combination of spatial separation (with one data series sent to the left stereo channel and one sent to the right stereo channel) and instruments of different timbres to present two data series in the same auditory graph. In a matching task, they found that graphs with two data series were generally more difficult to pair with their visual counterpart.

Brown and Brewster [66] later found that using different instruments for two data series had no effect on performance for drawing an accurate visual depiction of the auditory graph. They have, however, suggested that spatial separation should be employed when presenting multiple data streams [16]. Brown, Brewster, Ramloll, Yu, and Riedel [67] have also looked at concurrent (parallel) versus sequential (serial) presentation of multiple data streams and found that the best presentation mode may be dependent upon both task and user preferences, with parallel presentation mode perhaps being a preferable default [16]. Flowers [15], however, has emphasized that sequential presentations may be desirable for some tasks. He suggested that the concurrent display of numerous quantitative variables within the same auditory graph has generally not been effective, and he emphasized the importance of using distinct timbres when displaying multiple variables concurrently with sound. Issues of selective and divided attention that have driven much of the work on attention in the auditory modality [see 68] will likely be of great importance as multiple data streams are presented in auditory graphs.

3.4. Interaction and Mutual Influences of Listener, Task, and Display

For the purposes of organizing our discussion, we have attempted to place the variables contributing to auditory graph comprehension into three major groupings centered around the listener, the task, and the auditory graph. In practice, however, these major groups of variables do not compartmentalize well, as aspects related to the listener, the task, and the display interact in ways that are complex and not well-understood. Indeed, any discussion of a given group of variables outlined in this paper (e.g., listener variables) cannot proceed without crossing over to mention aspects of other variables (such as how the characteristics of the listener may influence the best design choices for a display, etc.). Current research [69, 70] has only begun to examine perceptual interactions that can result from manipulations of basic sound properties like frequency and amplitude. The crux of future research in auditory graph comprehension will be to investigate and come to understand the workings of these intricate relationships, interactions, and mutual influences of a broad range of relevant variables.

3.5. Environmental Considerations

The deployment of auditory graphs into ecologically valid scenarios may introduce unique environmental constraints and considerations. Some have suggested that the strict control exerted over stimuli and environmental conditions in most empirical research on auditory displays may cloud the generalizability of lab data to real world applications [48, 71]. A consideration of these issues is beyond the scope of this paper, and we have articulated our discussion under the assumption of more or less ideal listening conditions without environmental distractions, maskers, etc. As with any display, the real world application environment will be an important consideration, and innovative approaches to the software and hardware used to build and present auditory graphs will likely be of central importance to the success of auditory graphs in ecologically valid scenarios.

4. THE DATA EXPLORATION PROCESS

The data exploration process with auditory graphs, then, begins with a listener who possesses: 1) certain lawful processes of sensation and perception (e.g., thresholds of detection, susceptibility to forming auditory gestalts, etc.) that are here assumed to be more or less universal or similar across listeners; and 2) certain meaningful individual differences, which (although not well explained by current empirical data) should impact auditory graph comprehension. A listener with more musical experience, more domain expertise, better spatial ability, and/or better working memory ability should, all other things being equal, perform better on a task with auditory graphs than a listener with lower levels of these hypothesized relevant abilities.

The listener has a task, and the task dictates the information that the listener needs to extract from the display. When a person listens to an auditory graph, some data features may be perceived more or less automatically (i.e., as Gestalts or emergent features). If the emergent features contain the information required of the task, then the comprehension of the auditory graph should proceed with little effort or error. Performance measures, in these instances, should be near ceiling in most instances, and the impact of aforementioned individual differences may be negligible (i.e., all listeners should perform well regardless of individuals differences when the relevant percept emerges automatically).

In instances where emergent features are absent or do not contain the information necessary to fulfill the listener's task, the listener proceeds to more effortful processing of the auditory graph, and this subsequent extraction of information from an auditory graph likely proceeds in an iterative fashion. Kosslyn [7] suggested that task knowledge could prompt a person to "consciously *reorganize* the pattern" (p. 192) when the desired information is not available in the initial percept; for auditory graphs, we similarly suggest that listeners can actively manipulate their acquired representation of the data in an attempt to fulfill the needs of their particular tasks[2]. Others have similarly emphasized cyclical or iterative processes in graph comprehension that rely increasingly upon cognitive resources when the task requires information that is not available from (more or less automatic) perceptual processes [22, 72, 73]. Such effortful processing is necessarily more error prone than instances where the desired information was emergent in the display.

Certain individual difference variables (e.g., extensive training) may make some data features emergent (or at least allow for easier extraction of the desired information) for some listeners and not others. Likewise, design decisions regarding the display (e.g., mappings, scalings, polarities, and auditory context) will improve comprehension of the auditory graph to the extent that design decisions ease or accelerate the extraction of the information required for successful completion of the task. Future research and theory will help to specify more precisely the combinations of these variables that result in better or worse performance with auditory graphs, but it should be noted that

---

[2] With currently available data, it remains unclear the extent to which such reorganization or extended processing of the graph may be aided by repeated listening when the data are known and amenable to multiple presentations. It may be difficult or impossible to listen to the data more than once with auditory graphs of real time data, and this may be detrimental to tasks that require extended cognitive processing and reorganization of the initial percept.

auditory graph comprehension requiring attentional and cognitive resources should always suffer relative to comprehension of information that can be extracted automatically from the display.

5. CONCLUSIONS

We have presented an initial attempt to formulate a conceptual level model of auditory graph comprehension. We argue that research and application for auditory graphs should consider three broad categories of variables involving the listener, the task, and the auditory graph display. We recognize that our model is tentative and only one of perhaps many plausible descriptions of auditory graph comprehension, but we feel that the framework presented here offers a step forward toward a cohesive theoretical account of how humans interact with auditory graphs. We have articulated our arguments at a very high level, and in many or most cases we have intentionally avoided specifying the precise impacts of and relationships between variables and groups of variables. Clearly, a wealth of further research will be required to clarify and improve upon the ideas we present here. Our intention has been to provide a framework in which to organize the available data regarding auditory graph comprehension and to inspire future work that will expand upon the current knowledge and refine our understanding of how a listener comes to understand an auditory graph.

6. REFERENCES

[1] G. Kramer, B. N. Walker, T. Bonebright, P. Cook, J. Flowers, N. Miner, et al. "The Sonification Report: Status of the Field and Research Agenda. Report prepared for the National Science Foundation by members of the International Community for Auditory Display," International Community for Auditory Display (ICAD), Santa Fe, NM 1999.

[2] J. Edworthy, "Does sound help us to work better with machines? A commentary on Rautenberg's paper 'About the importance of auditory alarms during the operation of a plant simulator'," *Interacting with Computers*, vol. 10, pp. 401-409, 1998.

[3] A. Pirhonen, E. Murphy, g. McAllister, and W. Yu, "Non-speech sounds as elements of a use scenario: A semiotic perspective," presented at 12th International Conference on Auditory Display (ICAD06), London, UK, 2006.

[4] C. Frauenberger and T. Stockman, "Patterns in auditory menu design," presented at 12th International Conference on Auditory Display (ICAD06), London, UK, 2006.

[5] A. de Campo, "A data sonification design space map," presented at 2nd International Workshop on Interactive Sonification, York, UK, 2007.

[6] J. Zacks, E. Levy, B. Tversky, and D. Schiano, "Graphs in print," in *Diagrammatic Representation and Reasoning*, M. Anderson, B. Meyer, and P. Olivier, Eds. London: Springer, 2002, pp. 187-206.

[7] S. M. Kosslyn, "Understanding charts and graphs.," *Applied Cognitive Psychology*, vol. 3, pp. 185-225, 1989.

[8] S. Brewster, "Overcoming the lack of screen space on mobile computers," *Personal and Ubiquitous Computing*, vol. 6, pp. 188-205, 2002.

[9] S. Brewster and R. Murray, "Presenting dynamic information on mobile computers," *Personal Technologies*, vol. 4, pp. 209-212, 2000.

[10] S. Brewster, "Using non-speech sound to overcome information overload," *Displays*, vol. 17, pp. 179-189, 1997.

[11] M. L. Brown, S. L. Newsome, and E. P. Glinert, "An experiment into the use of auditory cues to reduce visual workload," presented at ACM CHI 89 Human Factors in Computing Systems Conference (CHI 89), 1989.

[12] B. N. Walker, "Magnitude estimation of conceptual data dimensions for use in sonification.," in *Journal of Experimental Psychology: Applied*, vol. 8: American Psychological Assn, 2002, pp. 211-221.

[13] R. N. Shepard, "Structural representation of musical pitch," in *The Psychology of Music*, D. Deutsch, Ed. New York Academic Press, 1982.

[14] M. Kubovy, "Concurrent pitch segregation and the theory of indispensable attributes," in *Perceptual Organization*, M. Kubovy and J. Pomerantz, Eds. Hillsdale, NJ: Lawrence Erlbaum, 1981, pp. 55-99.

[15] J. H. Flowers, "Thirteen years of reflection on auditory graphing: Promises, pitfalls, and potential new directions," presented at International Conference on Auditory Display (ICAD 2005), Limerick, Ireland, 2005.

[16] L. M. Brown, S. A. Brewster, R. Ramloll, M. Burton, and B. Riedel, "Design guidelines for audio presentation of graphs and tables," presented at International Conference on Auditory Display (ICAD2003), Boston, MA, 2003.

[17] T. Stockman, L. V. Nickerson, and G. Hind, "Auditory graphs: A summary of current experience and towards a research agenda," presented at International Conference on Auditory Display (ICAD2005), Limerick, Ireland, 2005.

[18] J. H. Flowers, D. C. Buhman, and K. D. Turnage, "Data sonification from the desktop: Should sound be part of standard data analysis software?," *ACM Transactions on Applied Perception*, vol. 2, pp. 467-472, 2005.

[19] B. N. Walker and M. A. Nees, "An agenda for research and development of multimodal graphs," presented at International Conference on Auditory Display (ICAD2005), Limerick, Ireland, 2005.

[20] S. Pinker, "A theory of graph comprehension.," in *Artificial Intelligence and the Future of Testing.*, R. Freedle, Ed. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc, 1990, pp. 73-126.

[21] C. M. Carswell, "Reading graphs: Interactions of processing requirements and stimulus structure.," in *Percepts, Concepts and Categories: The Representation and Processing of Information.*, B. Burns, Ed. North-Holland: Elsevier, 1992, pp. 605-645.

[22] P. A. Carpenter and P. Shah, "A model of the perceptual and conceptual processes in graph comprehension," *Journal of Experimental Psychology: Applied*, vol. 4, pp. 75-100, 1998.

[23] D. J. Gillan and R. Lewis, "A componential model of human interaction with graphs: I. Linear regression modeling.," *Human Factors*, vol. 36, pp. 419-440, 1994.

[24] C. M. Carswell, "Choosing specifiers: An evaluation of the basic tasks model of graph comprehension," *Human Factors*, vol. 34, pp. 535-554, 1992.

[25] G. Kramer, "An introduction to auditory display," in *Auditory Display: Sonification, Audification, and Auditory Interfaces*, G. Kramer, Ed. Reading, MA: Addison Wesley, 1994, pp. 1-78.

[26] J. H. Flowers, L. E. Whitwer, D. C. Grafel, and C. A. Kotan, "Sonification of daily weather records: Issues of perception, attention, and memory in design choices," presented at 2001 International Conference on Auditory Display (ICAD01), Espoo, Finland, 2001.

[27] K. D. Turnage, T. L. Bonebright, D. C. Buhman, and J. H. Flowers, "The effects of task demands on the equivalence of visual and auditory representations of periodic numerical data.," *Behavior Research Methods, Instruments & Computers*, vol. 28, pp. 270-274, 1996.

[28] H. Luczak, "Task analysis," in *Handbook of Hman Factors and Ergonomics*, G. Salvendy, Ed., 2nd ed. New York: Wiley, 1997, pp. 340-416.

[29] D. R. Smith and B. N. Walker, "Effects of auditory context cues and training on performance of a point estimation sonification task," *Applied Cognitive Psychology*, vol. 19, pp. 1065-1087, 2005.

[30] S. Barrass, "Auditory information design," Australian National University, 1997.

[31] W. S. Cleveland and R. McGill, "Graphical perception: Theory, experimentation, and application to the development of graphical

methods," *Journal of the American Statistical Association*, vol. 79, pp. 531-554, 1984.

[32] M. R. Jones, "Time, our lost dimension: Toward a new theory of perception, attention, and memory," *Psychological Review*, vol. 83, pp. 323-55, 1976.

[33] M. A. Nees, "Data density and trend reversals in auditory graphs: Effects on point estimation and trend identification tasks," Unpublished Master's Thesis: Georgia Institute of Technology 2007.

[34] A. Sandor and D. M. Lane, "Sonification of absolute values with single and multiple dimensions," presented at 2003 International Conference on Auditory Display (ICAD03), Boston, MA, 2003.

[35] B. N. Walker and M. A. Nees, "Brief training for performance of a point estimation task sonification task," presented at International Conference on Auditory Display (ICAD2005), Limerick, Ireland, 2005.

[36] M. A. Nees and B. N. Walker, "Relative intensity of auditory context for auditory graph design," presented at Twelfth International Conference on Auditory Display (ICAD06), London, UK, 2006.

[37] W. M. Hartmann, *Sounds, signals, and sensation: Modern acoustics and signal processing*. New York: Springer Verlag, 1997.

[38] T. D. Rossing, *The Science of Sound*. Reading, MA: Addison-Wesley, 1982.

[39] P. M. Sanderson, J. M. Flach, M. A. Buttigieg, and E. J. Casey, "Object displays do not always support better integrated task performance," *Human Factors*, vol. 31, pp. 183-198, 1989.

[40] A. Treisman and R. Paterson, "Emergent features, attention, and object perception," *Journal of Experimental Psychology: Human Perception & Performance*, vol. 10, pp. 12-31, 1984.

[41] M. Wertheimer, "Laws of organization in perceptual forms," in *A sourcebook of Gestalt psychology*, W. Ellis, Ed. London: Routledge & Kegan Paul, 1923/1938.

[42] S. M. Williams, "Perceptual principles in sound grouping," in *Auditory Display*, G. Kramer, Ed.: Addison-Wesley, 1994, pp. 95-125.

[43] A. S. Bregman, "Auditory scene analysis: Hearing in complex environments. ," in *Thinking in Sound: The Cognitive Psychology of Human Audition*, S. McAdams and E. Brigand, Eds. New York: Clarendon Press/Oxford University Press, 1993, pp. 10-36.

[44] A. S. Bregman, *Auditory Scene Analysis: The Perceptual Organization of Sound*. Cambridge, MA: MIT Press, 1990.

[45] D. Deutsch and J. Feroe, "The internal representation of pitch sequences in tonal music," *Psychological Review*, vol. 88, pp. 503-522, 1981.

[46] P. Vickers and B. Hogg, "Sonification abstraite/sonification concrete: An 'aesthetic persepctive space' for classifying auditory displays in the ars musica domain," presented at International Conference on Auditory Display (ICAD2006), London, UK, 2006.

[47] J. G. Neuhoff and J. Wayand, "Pitch change, sonification, and musical expertise: Which way is up?," presented at International Conference on Auditory Display, Kyoto, Japan, 2002.

[48] C. S. Watson and G. R. Kidd, "Factors in the design of effective auditory displays," presented at International Conference on Auditory Display (ICAD1994), Sante Fe, NM, 1994.

[49] E. G. Schellenberg and S. E. Trehub, "Good pitch memory is widespread," *Psychological Science*, vol. 14, pp. 262-266, 2003.

[50] B. N. Walker and L. M. Mauney, "Individual differences, cognitive abilities, and the interpretation of auditory graphs.," presented at International Conference on Auditory Display (ICAD2004), Sydney, Australia, 2004.

[51] S. B. Trickett and J. G. Trafton, "Toward a comprehensive model of graph comprehension: Making the case for spatial cognition," presented at Fourth International Conference on the Theory and Application of Diagrams (DIAGRAMS 2006), Stanford University, USA, 2006.

[52] E. Rusconi, B. Kwan, B. L. Giordano, C. Umilta, and B. Butterworth, "Spatial representation of pitch height: the SMARC effect," *Cognition*, vol. 99, pp. 113-129, 2006.

[53] J. A. Toth and C. M. Lewis, "The role of representation and working memory in diagrammatic reasoning and decision making," in *Diagrammatic Representation and Reasoning*, M. Anderson, B. Meyer, and P. Olivier, Eds. New York: Springer, 2002, pp. 207-221.

[54] J. Anderson, "Creating an empirical framework for sonification design," presented at International Conference on Auditory Display (ICAD2005), Limerick, Ireland, 2005.

[55] B. N. Walker and D. M. Lane, "Psychophysical scaling of sonification mappings: A comparision of visually impaired and sighted listeners," presented at 7th International Conference on Auditory Display, Espoo, Finland, 2001.

[56] E. Childs, "Auditory graphs of real-time data," presented at International Conference on Auditory Display (ICAD 2005), Limerick, Ireland, 2005.

[57] B. N. Walker, "Consistency of magnitude estimations with conceptual data dimensions used for sonification," *Journal of Applied Cognitive Psychology*, in press.

[58] P. Roth, H. Kamel, L. Petrucci, and T. Pun, "A comparison of three nonvisual methods for presenting scientific graphs," *Journal of Visual Impairment & Blindness*, vol. 96, pp. 420-428, 2002.

[59] S. A. Brewster, P. C. Wright, and A. D. N. Edwards, "An evaluation of earcons for use in auditory human-computer interfaces," presented at SIGCHI Conference on Human Factors in Computing Systems, Amsterdam, 1993.

[60] D. R. Smith and B. N. Walker, "Tick-marks, axes, and labels: The effects of adding context to auditory graphs," presented at International Conference on Auditory Display (ICAD2002), Kyoto, Japan, 2002.

[61] 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.," presented at International Conference on Auditory Display (ICAD2001), Espoo, Finland, 2001.

[62] W. W. Turnbull, "Pitch discrimination as a function of tonal duration," *Journal of Experimental Psychology*, vol. 34, pp. 302-316, 1944.

[63] J. M. Doughty and W. R. Garner, "Pitch characteristics of short tones II: Pitch as a function of tonal duration.," *Journal of Experimental Psychology*, vol. 38, pp. 478-494, 1948.

[64] W. R. Garner, "The accuracy of counting repeated short tones," *Journal of Experimental Psychology*, vol. 41, pp. 310-316, 1951.

[65] P. G. Cheatham and C. T. White, "Temporal numerosity: III. Auditory perception of number," *Journal of Experimental Psychology*, vol. 47, pp. 425-428, 1954.

[66] L. M. Brown and S. A. Brewster, "Drawing by ear: Interpreting sonified line graphs," presented at International Conference on Auditory Display (ICAD2003), Boston, MA., 2003.

[67] L. M. Brown, S. Brewster, and B. Riedel, "Browsing modes for exploring sonified line graphs," presented at 16th Britich HCI Conference, London, UK, 2002.

[68] N. Moray, *Listening and Attention*. Baltimore, MD: Penguin Books, 1969.

[69] J. G. Neuhoff, G. Kramer, and J. Wayand, "Pitch and loudness interact in auditory displays: Can the data get lost in the map?," *Journal of Experimental Psychology: Applied*, vol. 8, pp. 17-25, 2002.

[70] B. N. Walker and A. Ehrenstein, "Pitch and pitch change interact in auditory displays," *Journal of Experimental Psychology: Applied*, vol. 6, pp. 15-30, 2000.

[71] J. Neuhoff, "Ecological psychoacoustics: Introduction and history," in *Ecological Psychoacoustics*, vol. 1-13, J. Neuhoff, Ed. New York: Academic Press, 2004.

[72] P. Shah, "Graph comprehension: The role of format, content, and individual differences," in *Diagrammatic Representation and Reasoning*, M. Anderson, B. Meyer, and P. Olivier, Eds. New York: Springer, 2002, pp. 173-185.

[73] J. G. Trafton and S. B. Trickett, "A new model of graph and visualization usage," presented at 23rd Annual Conference of the Cognitive Science Society, 2001.