ONE SMALL STEP: SOUND SOURCES AND EVENTS AS THE BASIS FOR AUDITORY GRAPHS

John G. Neuhoff

The College of Wooster
Department of Psychology
Wooster, OH 44691
USA
jneuhoff@wooster.edu

Laurie M. Heller

Brown University
Department of Psychology
Providence, RI 02912
USA
Laurie_Heller@brown.edu

ABSTRACT

An overwhelming majority of auditory graphs employ a representational design that maps changes in a variable to changes in a “low-level” acoustic dimension such as frequency, intensity, or spectrum. However, there are several potential drawbacks to this type of auditory graph design. First, the perceptual correlates of these dimensions (pitch, loudness, and timbre) have been shown to interact perceptually such that changes in one dimension can influence judgments about changes in the others. Second, abstract changes in acoustic dimensions typically fail to invoke any kind of mental model that might help the listener represent cognitively the changes that occur in the data. Finally, listeners often are much better at attending to acoustic *sources* (the objects producing the sound) and acoustic *events* (the actions of these sounding objects) than to the low-level acoustic dimensions themselves. In this paper we endorse an approach to mapping data to sound that ties acoustic parameters unambiguously to changes in sound source or event characteristics. This type of design might be achieved by changing *complex* acoustic features along one axis in a manner that corresponds with a basic physical feature of a sound source or event.

1. REPRESENTING DATA WITH SOUND

The overwhelming majority of sonification techniques have employed a mapping scheme in which a variable that is to be sonified is mapped to an acoustic dimension such as frequency, intensity or spectrum. In some respects the utility of this type of representation seems entirely reasonable. Low-level acoustic dimensions are easily specified and can be easily mapped on to a wide variety of numerical data sets. The range of domains in which this approach been used is extremely diverse. For example, changes in frequency, intensity, and spectrum have been used to represent data from geological gas and oil explorations (Barrass & Zehner, 2000), anesthesiology and factory production controls (Fitch & Kramer, 1994; Gaver, Smith, & O'Shea, 1991), seismological investigations (Hayward, 1994; Saue & Fjeld, 1997; Speeth, 1961), and even the structural integrity of large bridges (Valenzuela, Sansalone, Krumhansl, & Streett, 1997). This general approach to the mapping problem in sonification has been mirrored to a large extent in the more specific domain of auditory graph design.

1.1. Mapping Approaches in Auditory Graphs

Perhaps the most widely used dimension in the design of auditory graphs is acoustic frequency. In fact, it is rare to encounter an auditory graph that does *not* employ some type of representational frequency change. Among other things, frequency change has been used to represent live birth rates, human developmental biometrics, and insect mortality (Bonebright, Nees, Connerley, & McCain, 2001). Frequency change has also been used to graphically display weather patterns (Flowers & Grafel, 2002), and to investigate the ability of listeners to draw generic line graphs based on their auditory presentation (Brown & Brewster, 2003).

In addition to frequency based auditory graphs in which changes in frequency over time represent changes in a data variable, some designers have combined multiple acoustic cues in order to represent more than one variable. Neuhoff, Kramer & Wayand (2002), used changes in pitch and loudness to represent changes in the price and trading volume of stock market data respectively. Barrass & Zehner (2000) used changes in pitch, timbre, and repetition rate to represent various characteristics of geological data.

However, the practice of using low-level acoustic dimensions to represent data in an auditory graph can present difficulties for the listener on several fronts. First, auditory dimensions such as pitch, loudness, and timbre have been shown to interact perceptually. A substantial body of work has shown that changes in one auditory dimension can influence judgments about changes in the others (Garner, 1974; Grau & Kemler-Nelson, 1988; Melara & Marks, 1990; Melara, Marks, & Potts, 1993; Neuhoff & McBeath, 1996; McBeath & Neuhoff, 2002; Neuhoff, McBeath, & Wanzie, 1999; Shepard, 1964). This lack of dimensional orthogonality presents a particular challenge to designers of auditory graphs who simultaneously represent changes in more than one data set with changes in more than one auditory dimension. The difficulty arises because acoustically represented changes in one variable can cause unintended changes to be perceived in the other variables (Neuhoff, Kramer & Wayand, 2002).

Second, abstract changes in low-level auditory dimensions typically fail to invoke any kind of mental model that might help the listener represent cognitively the changes in the data (Kramer, 1995). Changes in pitch or loudness, for example, can be the result a number of physical changes in the characteristics of the source that produces the sound. Under natural listening conditions, changes in loudness could occur because of changes in the intensity of the sound produced by the source, because of changes in the distance between the source and the listener, or because of an occluding object that moves between the source and the listener. Thus, the listener has no concrete model on which the abstract low-level acoustic change can be based.

Finally, listeners often are much better at attending to acoustic *sources* (the objects producing the sound) and acoustic *events* (the actions of these sounding objects) than to the auditory dimensions that give rise to the impression of the source. Gaver (1993) has suggested that in everyday listening, listeners focus their attention on the identity of the source that produces the sound and on the physical characteristics of that source. These three fundamental challenges to the use of low-level acoustic dimensions in auditory graph design--dimensional interaction, absence of mental models, and attention to sound source characteristics, are outlined in more detail below.

1.2. Interacting Perceptual Dimensions

Previous research has shown that when listeners evaluate an auditory stimulus on a dimension of interest (such as pitch), they sometimes encounter interference from orthogonal variation of an unattended dimension (e.g., loudness). The two dimensions interact and are considered "integral" (Garner 1974). According to one view, the individual dimensions constituting the stimuli in effect are not perceived directly, but are processed in a holistic manner (Shepard, 1964). In other words the listener does not have primary access to dimensions in question, and cannot selectively attend to one dimension. Thus, the concurrent use of interacting perceptual dimensions to represent separate variables can be problematic because listeners are unable to attend to individual variables selectively. This is not to say that a dimensional structure cannot be extracted from integral dimensions, but that it is a more derived and secondary cognitive process (Garner 1974; Kemler-Nelson, 1993).

Alternatively, Melara and Marks (1990; Melara, Marks, & Potts, 1993) have advanced a model of dimensional interaction that proposes that the extraction of a dimensional attribute creates a context in which attributes of the other dimension are perceived. In the case of interacting dimensions then, the perception of an attribute on one dimension is influenced by the context created by an attribute in the other dimension. In the words of Melara and Marks, "the attribute high pitch has one perceptual meaning when paired with the attribute loud but a different meaning when paired with the attribute soft. Context established by loudness values thus acts to weight perceptually the extraction of pitch information…" (Melara & Marks, 1990, p.399.)

Unfortunately, for the purpose of auditory graph design, the end result of interacting perceptual dimensions is the same regardless of the specifics of the underlying perceptual mechanism. The consistent finding that auditory dimensions interact presents a tremendous challenge for designers of auditory graphs. Perceptual interaction can distort changes in graphical data that are represented by low-level acoustic attributes.

1.3. Mental Models and Acoustic Change

Simple changes in pitch, loudness and timbre do not typically invoke a strong mental model of a specific acoustic object or event. This ambiguity might fuel further confusion in the interpretation of an auditory graph. If listeners had a strong auditory mental model they might better interpret auditory graphs in the context of the model. In other words, the mental model might provide some pre-existing cognitive structure in which to interpret the sonified changes in a variable.

Perhaps we should try turning our goals inside out. Rather than creating smooth changes in simple acoustic attributes, perhaps we should create smooth changes in mental models of sound sources. What does it mean to create a mental model? The sound should convey a sound source that the listener is already familiar with in the physical world, and that sound source should naturally vary its acoustics in a manner that maps to a single cognitive dimension. For example, if a listener hears the sound of an engine, they can build a mental model of a vehicle going at a certain speed; when the engine sound changes, it should either increase or decrease the perceived speed of that vehicle. Changes in perceived speed could map linearly to the desired data axis. Frequency and amplitude may not change linearly as speed increases. Instead, perceptual tests would have to establish what acoustic changes are necessary to create a smooth and evenly spaced increase in perceived vehicle speed. This could define one axis of the graph even though the acoustic changes along this axis might be complex, unequal, or difficult to describe. This "event-based" approach to auditory display was first suggested by Kramer (1994). The idea was to use a "virtual engine", (a theoretical sonification using an automobile engine sound) as a display metaphor for a data set unrelated to physical engines.

This kind of mapping would take advantage of knowledge that the listener already has about how sound sources behave, and how their acoustics should change. The listener's mental model of a physical event should do the work to convey the variations in the data. This should reduce the mental effort required for translation from the acoustic display to the conceptual understanding of the data. An interface that uses mental models should be more intuitive, easier to learn, and faster to use.

1.4. Attending to Auditory Events

Gaver (1993) has argued that listeners do not typically listen to, and in some cases do not even easily identify, changes in pitch, loudness or timbre. Instead he proposed that listeners attend to acoustic events and sources in the environment. Changes in a sound are not typically characterized as changes in the acoustic attributes of the signal, but rather as physical changes in the properties or dynamic characteristics of the source. In other words, everyday listeners would be less likely to describe a sound as "a high intensity narrow band transient with a fast decay, a spectral centroid of 4khz, and a 2 Hz repetition rate" and more likely to describe it as "someone hammering a nail".

It may be then that a mapping technique that employs acoustic events or source characteristics might make it easier for listeners to attend to the relevant changes in an auditory graph. Such an approach might also more effective in overcoming the other challenges presented by the use of low-level acoustic dimensions.

2. MAPPING DATA CHANGE TO AUDITORY EVENTS

There is a strong theoretical basis for connecting auditory perception with sources and events rather than with acoustic dimensions. Sound is generated by the physical interactions of objects, surfaces, and substances. The sound waveform can carry a great deal of potential information about its source's properties. However, the acoustic attributes of sources and events are complex and time-varying; no single acoustic attribute specifies a particular property of an object. In fact, the situation is worse than that. The same single acoustic parameter (e.g. frequency) can be affected by multiple attributes of the source (e.g. density, homogeneity, shape, size...). This means that as one acoustic feature such as frequency is varied, there are multiple possible interpretations for what is happening in the physical world to cause that change frequency. This uncertainty may impose a cognitive load when interpreting graphing structures that vary frequency as an information-bearing dimension.

We suggest that it is possible to find changes in acoustic parameters that map unambiguously to changes in source or event characteristics. This might be achieved by changing *complex* acoustic features along one axis in a manner that corresponds with a basic physical feature or even associated with a sound source. For example, one could indicate a source's solidity along a continuum from solid to liquid, with muddy material in the middle of the range. Of course, this approach would vary more than a single frequency at once, but these are precisely the kinds of complex dynamic signals that the auditory system is equipped to deal with. To add a second axis, one might vary the attributes of the action (instead of the material). For example, the speed of footsteps could be varied, a salient stimulus that human listeners have been show to be very sensitive too (Li, Logan, & Pastore, 1991). A point on the lower left of the graph might be slow walking in water; a point on the upper right might be fast walking on concrete. As you move around this graphing space, the source/event would vary in a continuous way that would hopefully give rise to intuitive perceptual interpretations. This perceptual simplicity would be achieved at the cost of adding acoustic complexity to the graphing dimensions.

3. REFERENCES

Barrass, S., & Zehner, B. (2000). Responsive Sonification of Well Logs. *Paper presented at the International Conference on Auditory Display*, Atlanta, G. A.

Bonebright, T., Nees, M., Connerley, T. and McCain, G. (2001). Testing the effectiveness of sonified graphs for education: A programmatic research project", in *Proceedings of the International Conference on Auditory Display,* Espoo, Finland,

Brown, L.M. and Brewster, S.A., (2003). Drawing by Ear: Interpreting Sonified Line Graphs. *Proceedings of the International Conference on Auditory Display,* Boston, MA 152-156.

Fitch, T., & Kramer, G. (1994). Sonifying the body electric: Superiority of an auditory over a visual display in a complex, multi-variate system. In G. Kramer (Ed.), *Auditory display: Sonification, audification and auditory interfaces. Proceedings of the First International Conference on Auditory Display* (pp. 307-326). Reading, MA: Addison-Wesley.

Flowers, J. H. & Grafel, D. C. (2002). Perception of sonified daily weather records. *Proceedings of the Human Factors and Ergonomic Society 46th Annual Meeting,* 1579-1583.

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

Gaver, W. W. (1993). What in the world do we hear? An ecological approach to auditory event perception. *Ecological Psychology*, 5 (1), 1-29.

Gaver, W. W., Smith, R. B., & O'Shea, T. (1991). Effective sounds in complex systems: the ARKola simulation. *Paper presented at the CHI, New York.*

Grau, J. A., & Kemler-Nelson, D. G. (1988). The distinction between integral and separable dimensions: Evidence for the integrality of pitch and loudness. *Journal of Experimental Psychology: General, 117*, 347-370.

Hayward, C. (1994). Listening to the earth sing. In G. Kramer (Ed.), *Auditory display: Sonification, audification and auditory interfaces. Proceedings of the First International Conference on Auditory Display (ICAD)* (pp. 369-404). Reading, MA: Addison-Wesley.

Kramer, G. (1994b). Some organizing principles for auditory display. In G. Kramer (Ed.), *Auditory Display: Sonification, Audification, and Auditory Interface, SFI Studies in the Sciences of Complexity, Proc. XVIII*. Reading, MA: Addison-Wesley.

Kramer, G. (1995). Sound and communication in virtual reality. In F. Biocca & M. Levy (Eds.) *Communication In The Age Of Virtual Reality*, 259-276, Lawrence Earlbaum.

Kramer, G. (1995). Sound and Communication in Virtual Reality.In F. Biocca & M. Levy (Eds.) *Communication In The Age Of Virtual Reality*, Lawrence Earlbaum.

Li, X.-F., Logan, R.J., & Pastore, R.E. (1991). Perception of acoustic source characteristics: Walking sounds. *Journal of the Acoustical Society of America* , 90, 3036-3049.

McBeath, M. K., & Neuhoff, J.G. (2002). The Doppler effect is not what you think it is: Dramatic pitch change due to dynamic intensity change. *Psychonomic Bulletin & Review 9 (2)* 306-313.

Melara, R. D., & Marks, L. E. (1990). Perceptual primacy of dimensions: Support for a model of dimensional interaction. *Journal of Experimental Psychology Human Perception and Performance, 16*, 398-414.

Melara, R. D., Marks, L. E., & Potts, B. C. (1993). Early-holistic processing or dimensional similarity? *Journal of Experimental Psychology Human Perception and Performance, 19*, 1114-1120.

Neuhoff, J. G., & McBeath, M. K. (1996). The Doppler illusion: The influence of dynamic intensity change on perceived pitch. *Journal of Experimental Psychology: Human Perception and Performance, 22*, 970-985.

Neuhoff, J. G., McBeath, M. K., & Wanzie, W. C. (1999). Dynamic frequency change influences loudness perception: A central, analytic process. *Journal of Experimental Psychology: Human Perception and Performance, 25*(4), 1050-1059.

Neuhoff, J.G., Kramer, G., & Wayand, J. (2002). Pitch and loudness interact in auditory displays: Can the data get lost in the map? *Journal of Experimental Psychology: Applied.* 8 (1), 17-25.

Saue, S., & Fjeld, O. K. (1997). A Platform for Audiovisual Seismic Interpretation. *Paper presented at the International Conference on Auditory Display*, Palo Alto, CA.

Shepard, R. N. (1964). Attention and the metric structure of the stimulus space. *Journal of Mathematical Psychology, 1,* 54-87.

Speeth, S. D. (1961). Seismometer sounds. *Journal of the Acoustical Society of America, 33*, 909-916.

Valenzuela, M. L., Sansalone, M. J., Krumhansl, C. L., & Streett, W. B. (1997). Use sound for the interpretation of impact-echo signals. *Paper presented at the International Conference on Auditory Display*, Palo Alto, CA.