OCAD University Open Research Repository

Faculty of Design, Industrial Design

2015

What is lost in translation from visual graphics to text for accessibility

Coppin, Peter

Suggested citation:

Coppin, Peter (2015) What is lost in translation from visual graphics to text for accessibility. In: Proceedings of the EuroAsianPacific Joint Conference on Cognitive Science, 25-27 Sept 2015, Torino, Italy. Available at http://openresearch.ocadu.ca/id/eprint/1034/

What is Lost in Translation from Visual Graphics to Text for Accessibility

Peter Coppin (pcoppin@faculty.ocadu.ca)
Dept. of Industrial Design, Faculty of Design, OCAD University, Toronto, ON M5T 1W1 CANADA
Dept. of Mechanical and Industrial Engineering, University of Toronto, Toronto, ON M5S 3G8 CANADA

Abstract

Many blind and low-vision individuals are unable to access digital media visually. Currently, the solution to this accessibility problem is to produce text descriptions of visual graphics, which are then translated via text-to-speech screen reader technology. However, if a text description can accurately convey the meaning intended by an author of a visualization, then why did the author create the visualization in the first place? This essay critically examines this problem by comparing the so-called graphic–linguistic distinction to similar distinctions between the properties of sound and speech. It also presents a provisional model for identifying visual properties of graphics that are not conveyed via text-to-speech translations, with the goal of informing the design of more effective sonic translations of visual graphics.

Graphics Without Visual Perception

Consider the experience of a blind or low-vision individual who uses a screen reader to access pictures, diagrams, charts, and graphs. Unlike a user who accesses graphical media through visual perception, the screen reader user usually accesses these graphics via text-to-speech "descriptions," essentially interpretations of what was deemed most relevant by the person who produced the text descriptions of the author's intended meaning. For example, Figure 1a presents a financial chart with rising and falling stock prices over time, where time is shown on the horizontal axis and monetary value is shown on the vertical axis. Figure 1d presents a text description of the chart compliant with the Web Content Accessibility Guidelines (WCAG), using text to describe the rising and falling monetary values over time. The next sections compare and contrast how these presentations are experienced.

In a text description of a visual graphic (Figure 1d), all of the information is conveyed via text (or text-to-speech, when conveyed via screen reader technology). But in the original chart (Figure 1a), only some of the information is conveyed via text, predominantly numerical values and labels (Figure 1c); the shape of the shaded contour (Figure 1b) is not conveyed via text: the visually perceived shapes are picked up "more directly" and the features of shapes are translated to text descriptions. However, important properties of visually perceived shape information (Figure 1b) are lost in translation and are instead conveyed via text (Figure 1e). This shape information is needed to provide the unique affordances that are often associated with "visual" representations relative to text.
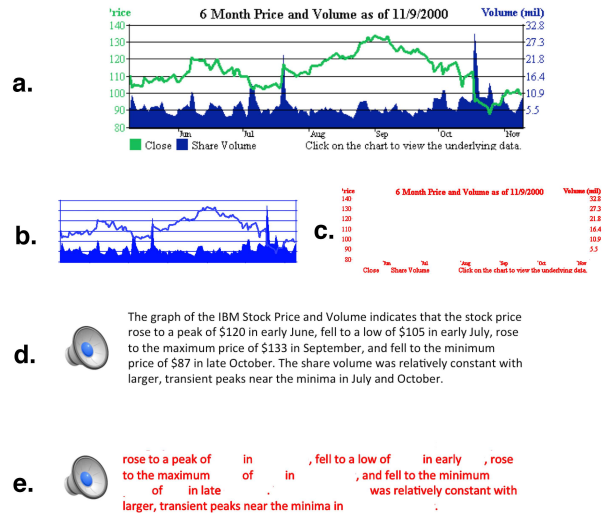

Figure 1. The chart (a) is composed of visually perceived shape contours (b) and text labels (c). Accessibility practices translate b–c to text (d), with shapes described via text (e).[1]

Many scholars have explored the differences between graphics and text, often referred to as the so-called "graphic–linguistic distinction" (Shimojima, 1999). In addition, researchers have investigated how so-called "non-linguistic sonification" can be employed to make charts and graphs more accessible (e.g., Hermann, Hunt, & Neuhoff, 2011; Mauney & Walker, 2004). This essay examines the graphic–linguistic distinction in order to better understand how it could correspond to a similar distinction between properties of non-linguistic sonification compared to speech to provide a means to identify what is lost when graphics are translated to text-to-speech. An increased understanding could inform the design of new approaches for conveying properties of graphically represented shapes via sound.

The Graphic–Linguistic Distinction: Implications for Sonic Interface Design

The graphic–linguistic distinction has been described in various ways: analogical versus Fregean; analog versus propositional; graphical versus sentential; and diagrammatical versus linguistic (Shimojima, 1999).

---

[1] Adapted from "Web Accessibility Best Practices: Graphs" by Campus Information Technologies and Educational Services (CITES) and Disability Resources and Educational Services (DRES), University of Illinois at Urbana/Champaign. Copyright 2005 by University of Illinois at Urbana/Champaign.

2D Versus Sequential

According to Larkin and Simon (1987), diagrammatic representations "preserve explicitly the information about the topographical and geometric relations among the components of the problem." More specifically, Larkin and Simon defined a diagram as "a data structure in which elements appear in a single sequence" whereas a diagrammatic representation is a "data structure in which information is indexed by two-dimensional location." For the purposes of this essay, the text description in Figure 1e is classified as sentential because the text is composed of marks arranged in a linear sequence and the marks are taken to refer to words with linguistic meanings (linguistically conveyed elements). In contrast, Figure 1a is classified as a diagram because the financial values are indicated via (textually) labeled points or lines (elements) that are indexed to a graphical grid. The visually processed spatial relations among these labeled marks yield powerful affordances, because by processing the contours of lines or the relative positions of marks scattered across the two-dimensional graphical surface, the viewer can infer values and trends that are not explicitly conveyed via labels (cf. Barwise & Etchemendy, 1990).

Implications for sonic charts and graphs

Sonic sentential properties. Text-to-speech (the current standard for WCAG accessibility) would seem to be the obvious candidate for the sonic version of what Larkin and Simon referred to as a sentential structure, where elements are arranged in a linear sequence. In the case of visually processed written sentences composed of word forms printed on a page, the sequential properties result from the linear arrangement of characters and word forms on the printed surface. In the case of sonic sentential structures, the sequential properties are temporal, presented as a sequence of sounds that are perceptually processed as words that refer to intended meanings. Larkin and Simon did not define what the elements (that are arranged in sequence) are composed of. For the purpose of this subsection, let us assume that the elements are some combination of properties that, when sequentially processed as words, refer to intended items.

Sonic diagrammatic properties. To present diagrammatic properties in a way that can be perceived aurally, designers would need to exploit properties of sound that can convey topological and geometric relations. People use stereo, echo, and the Doppler effect to determine the spatial locations of sound-producing objects in physical environments (cf. Nasir & Roberts, 2007). Designers could exploit these cues to convey geometric and topological relations among elements that are indexed to a 2D plane (cf. Brown, Brewster, Ramloll, Burton, & Riedel, 2003; Hermann, Hunt, & Neuhoff, 2011). Figure 2 shows how left and right arrow keys could move an "audio cursor" to different positions on an *x*-axis of a computationally generated 2D space. The position of the sonically conveyed cursor on the *x*-axis could be indicated via stereo (cf. Zhao, Plaisant, Shneiderman, & Lazar, 2008). For a simple spark line graph, the sonic cursor

can alter the pitch of the sound if "scrubbed" to different points on the *x*-axis, so that higher pitches correspond to points that intersect with the cursor at higher elevations (Figure 2, right) and lower pitches correspond to points that intersect with the cursor at lower elevations, thereby allowing blind or low-vision users to perceive the contours of the graph (cf. Brown, Brewster, Ramloll, Burton, & Riedel, 2003).

Relation Symbols and Object Symbols

According to Russell (1923), in sentences "words which mean relations are not themselves relations," whereas in graphical representations like maps, "a relation is represented by a relation." An example of the latter is the financial chart (e.g., Figure 1a), where higher monetary values are conveyed via marks at higher elevations of the graphic, whereas lower monetary values are conveyed via marks at lower elevations. This convention allows the visually perceived spatial relationships among the marks to represent relationships among monetary values over time.

Implications for sonic charts and graphs

Graphical relations could be conveyed sonically. Consider two tones with different pitches: Tone A and Tone B (Figure 2, right). If Tone A is at a lower frequency than Tone B, then the sonic relation between the two tones is the perceptible difference in pitch between the tones. For example, if Tone A refers to a stock price at an earlier point in time, and Tone B refers to a stock price at a later point in time, then the perceptible difference between the pitches of the tones can convey the difference in price over time. Moving the sonic cursor from left to right would correspond to a change (increase) in pitch, conveying the change in stock price over time via a sonic relation.
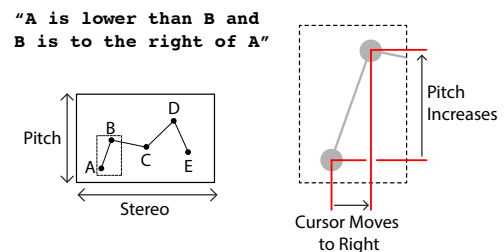

Figure 2. By scrubbing a "sonic cursor" along an axis, audiences could access sonically conveyed relations through changes in pitch and via stereo.

Analog Versus Digital

The classic distinction between analog versus digital, where *analog* refers to visual properties of a graphic and *digital* refers to linguistic properties, is most commonly associated with Goodman (1968). Shimojima (1999) illustrated this distinction using the example of a speedometer dial. The analog aspect of the dial is the perceived orientation of the needle relative to the numerically labeled marks on the dial. The digital aspect is the numerical magnitude (speed)

determined by comparing the needle's position to the marks representing numerical values.

Implications for sonic charts and graphs

The analog versus digital distinction appears to involve two interrelated capabilities: lower-level perceptual capabilities to process geometric and topological properties (e.g., those shown on the speedometer dial); and higher-level capabilities to process, filter, and interpret how those perceptually processed features fall into conceptual categories (e.g., the numerically represented velocity) (Mandler, 2006; Figure 3). A financial chart is composed of both analog (lines that vary along the Cartesian plane) and digital (text that conveys monetary values relative to time) properties. The current text-to-speech approach only exploits the digital properties of language (words with symbolic meaning) – but designers could produce effective visual-to-audio translations by recruiting features of sound such as pitch, echo, stereo, and timbre to convey properties of graphics that vary in analog space.
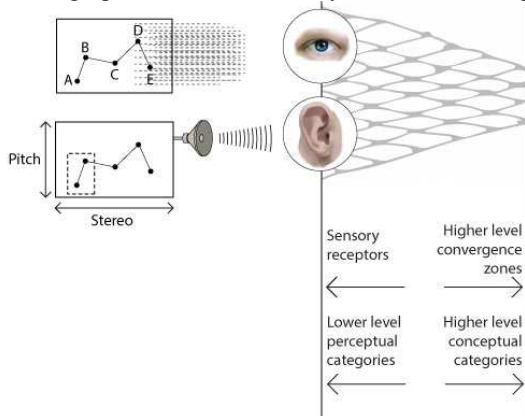

Figure 3. A perception–reaction system is hierarchically organized to process lower-level perceptual structures and categorize them into higher-level conceptual categories.

Intrinsic Versus Extrinsic Constraints

For brevity, the following discussion will use the classic characterization provided by Barwise and Etchemendy (1990) because it is compact and intuitive:

Diagrams are physical situations. They must be, since we can see them. As such, they obey their own set of constraints . . . By choosing a representational scheme appropriately, so that the constraints on the diagrams have a good match with the constraints on the described situation, the diagram can generate a lot of information that the user never need infer. Rather, the user can simply read off facts from the diagram as needed. This situation is in stark contrast to sentential inference, where even the most trivial consequence needs to be inferred explicitly.

To illustrate how "diagrams are physical situations," consider the illustration shown in Figure 2 (left). A text (or text-to-speech) description might go as follows: "A *is below* B and both A and B *are to the left of* C." Another textual description might read: "B *is between* A and C and *is above*

both A and C." Each text description conveys a different interpretation of what is shown visually and therefore affords different inferences. In contrast, a diagram can convey many other relationships because of how it conveys topological and geometric information through visual perception: Barwise and Etchemendy referred to this as a diagram's ability to present "countless facts."

In translating a visual diagram into a sonic diagram, one is essentially translating a spatial structure from one set of modality-specific perceptual features into another. Loomis et al. (2013) have argued that regardless of the stimulus modality, what is generated is ultimately an amodal "spatial image". Thus in principle it should be possible to convey the same underlying spatial representation with the unique affordances of diagrams using sonification.

Implications for sonic charts and graphs

When Barwise and Etchemendy (1990) referred to diagrams as "physical situations," they were referring to the properties (and affordances) of diagrams that emerge through interaction via a human visual perception system. The challenge for designers who seek to extend the affordances of visual diagrams to the sonic domain is to identify properties or dimensions of sound that similarly (i.e., using human perceptual processing of sound) make use of "physical situations" to present "countless facts."

Thus, a hybrid stereo–varying frequency interface (see Figure 3) should enable a user to "hear the shape" of a contour. Indexing text-to-speech labels to contours should allow users to form multiple sentences (countless facts) about the geometric and/or topological relations among the labeled elements.

Extending the Graphic–Linguistic Distinction into the Sonic Domain

Let us now extend on the various graphic–linguistic distinctions to consider sonic versions of visual charts and graphs.

1. Extending on the diagrammatic versus sentential distinction, text-to-speech can be considered a sonic version of what Larkin and Simon referred to as a sentential structure and is the current WCAG approach to web accessibility. In contrast, spatial sound can be exploited to convey 2D sonic diagrammatic external representations.

2. Extending on the analog versus digital distinction, text-to-speech uses language to convey digital properties sonically. The analog properties of sound, such as tone, timbre, stereo, and echo could afford the communication of spatial, geometric, or topological information.

3. Extending on the distinction between relation symbols and object symbols, the current text-to-speech approach uses words to convey relations. Because relations among elements represented by analog and spatial properties of sound are themselves relations, analog and spatial properties of sound could be recruited to map numerical values to perceptual dimensions.

4. Extending on the distinction between intrinsic and extrinsic constraints, producing sonic versions of visual graphics would require identifying "physical situations" that naturally emerge during human perceptual processing of sound to present "countless facts."

Perceptual and Conceptual Graphic Relations

This section integrates these extensions and proposes how the graphic–linguistic distinction could be extended to sonic external representations. First, let us recruit and expand on the distinction between lower-level perceptually processed topological and geometric features of an environment versus the recognition, categorization, and linguistic communication of those features.

Visual and aural sentential structures and relations are detected and perceptually processed via lower-level sensory receptors and perceptual categories (Figure 3, left). In written text or text-to-speech, what is most relevant is the higher-level conceptual category (Figure 3, right) that a given feature (such as perceptually processed printed text on a page or text-to-speech) is taken to fall under. What is needed is a way to convey topological and geometric relations among elements by exploiting lower-level perceptually processed features of a visual graphic or sonic structure (Figure 3, left). Let us refer to these perceptually processed features as *perceptual properties*. Let us refer to these perceptually processed relations among elements as *perceptual relations*. Let us refer to relations that are communicated via text as *text-described relations*.

Perceptual Relations vs. Text-Described Relations

We are now ready to build on previous work by Coppin (2014) to provide a theoretical foundation for distinguishing perceptual relations versus text-described relations.

The model is based on the idea that an individual's perception–reaction loop (cf. Gibson, 1986) enables survival and prosperity within a dynamic environment composed of change and variation. This requires capabilities to predict, anticipate, and *simulate* (Barsalou, 1999) dynamic change and variation. For example, reaching for and grasping an item such as a cup requires capabilities to perceptually process features from the proximal surface of the item and also to predict, anticipate, and simulate features of the distal surface of the item.

These simulations are constructed from the memory traces of past perception–reactions (conjunctive neurons), so simulation involves many of the same neural systems used during perception (Kosslyn, Ganis, & Thompson, 2001). For example, as I perceive the cup, I am also informing potential action (reaching for and grasping the proximal and distal sides of the cup). Thus, perception and simulation are integrated aspects of perception–reaction within a physical environment, and each act of perception–reaction leaves memory traces in the form of conjunctive neurons across lower-level association areas (Figure 3).

At lower-level association areas, which are more tightly coupled with sensory receptors, simulated prototypes fall under perceptual categories, At higher-level association areas (see Figure 3, right), conjunctive neurons converge in zones across multiple sensory modes. These "convergence zones" (Damasio, 1989; Simmons & Barsalou, 2003) enable simulated prototypes of possible perception–reactions that are not as easily described in terms of a specific perceptual mode or a reenactment of a specific prior perception–action. Instead, these simulated prototypes fall under more general categories of possible perception–actions (Barsalou, 2003). These are not only more amodal, but have been described as more filtered, interpreted (Pylyshyn, 1973), conceptual (Barsalou, 2003, 2005), or abstract (Barsalou, 2003). For example, a child who takes a bite out of what turns out to be a rotten apple might later reenact this experience when she perceives another rotten apple with common properties. Over time, she will develop an understanding of 'rotten' as a category that can include apples, as well as many other objects and experiences.

Back to charts and graphs. In a financial chart (and many other kinds of diagrams), relations are conveyed via lower-level perceptual processing of the geometrical and topological properties of the marked physical surface (Table 1). In contrast, in text descriptions (sentential structures), relations are conceptual (and conveyed *linguistically*; see Table 2); although visual properties of printed text or aural properties of text-to-speech are also picked up by sensory receptors, what is meaningful about them is the conceptual relation that is conveyed linguistically.

Table 1. Diagrams are composed of perceptually processed relations among linguistically conveyed conceptual objects; sentences are composed of linguistically conveyed conceptual relations among linguistically conveyed conceptual objects (adapted from Coppin, 2014).

	Diagrammatic	Sentential
Relations	*Perceptual*	*Conceptual*
Objects or Items	*Conceptual*	*Conceptual*

Perceptual Specificity is Lost in Translation

The idea of "specificity" is central to understanding what is lost in translation, so let us begin by clarifying what is meant by "more or less specific" in this context. Consider the line shown in Figure 4b. Relative to the line of Figure 4c, we have more knowledge about the location of a point in a one-dimensional space, due to the shaded red marker. This means we have more certainty (or more information) about the specified location of the point in Figure 4b than we do about the location of the point in Figure 4c.
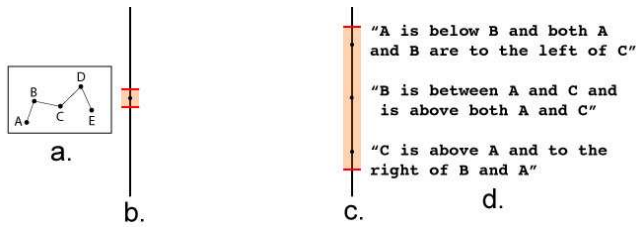
Figure 4. The left vertical line (b) refers to the limited range of perceptual structures conveyed via a given graphic. The right line (c) refers to the wider range of possible conceptual categories that the perceptual structures could fall under. The model predicts that when *perceptual* specificity is high (b) *conceptual* specificity is low (c).

Extending the line example to discuss perceptual relations, Figure 4b refers to intentionally configured marks or sounds from an author to cause intended audience percepts (the diagram in Figure 4a). However, the perceptual relations of Figure 4a can be processed, filtered, and interpreted to fall under a range of possible relational categories (that can be text-described), indicated by the highlighted segment of the right line in Figure 4c (as shown in Figure 4d: "A *is below* B and both A and B *are to the left of* C" or "B *is between* A and C and *is above both* A and C"). In other words, although perceptual specificity is high, conceptual specificity of the intended relation is low because the perceptual relations can fall under numerous conceptual categories. However, the reverse is also true and this reversal exposes the heart of what is lost during the translation process.

Conceptual Specificity is Perceptually Ambiguous
Extending the line example to discuss the perceptual ambiguity of text-described (conceptual) relations, the right highlighted line in Figure 5c refers to a specific (sentential) text description authored to convey intended conceptual relations (Figure 5d). However, numerous perceptual relations (Figure 5a) can fall under the text-described conceptual relations, indicated by the highlighted segment of the left line in Figure 5b. In other words, although conceptual specificity is high, perceptual specificity of the intended relations is low, because numerous perceptual relations can fall under the text-described conceptual relations.
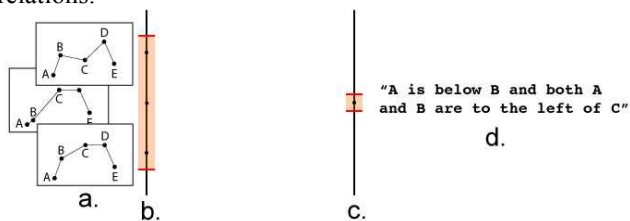

Figure 5. The model predicts that when *conceptual* specificity is high (c) *perceptual* specificity is low (b).

Application to an Example Design Problem

Let us now return to the WCAG text description example from Figure 1 in order to demonstrate what is lost in

translation and how what is lost could be conveyed via non-linguistic sound. In the text description (Figure 1d), the problem is that all content is conveyed conceptually (via text-to-speech) whereas the original visual graphic that the text description is based on conveys much of the content (the contour of the shape) perceptually: *Perceptual relations are lost and replaced by conceptual relations, generating perceptual ambiguity*. If the objective is to present Figure 1a sonically, how can a designer decide which aspects should be conveyed via conceptual properties (text-to-speech) and which aspects should be conveyed via perceptual sonic properties (such as spatial sound)?

Recall the perceptual distinction, where perceptual properties are predicted to afford the communication of concrete structures more effectively compared with conceptual properties, and an aspect of a graphic can be identified as "more concrete" if it produces a perceptual structure that corresponds to what could be picked up and perceptually processed from a physical environment. In this account, the graphically represented shape contour (Figure 1b) is primarily perceptual, and is therefore more appropriate for translation to sonic properties that can use spatial sound to convey geometric and topological relations among conceptually conveyed objects.

To determine which aspects of a graphic should be conveyed via text-to-speech, recall the conceptual distinction: text is predicted to afford the communication of abstract conceptual categories more effectively compared with perceptual properties, and a concept can be identified as more abstract if it is less tied to a specific perceptual mode). In other words, it is *less easily mapped back* to a structure that could be picked up and perceptually processed from a physical environment. Under this account, the numbers that label increments on the *x* and *y* axes (Figure 1a) are more conceptual because they cannot be mapped back to a perceptual structure that could be picked up from a physical environment.

Conclusion

This essay proposes a provisional model to underpin the various accounts of the graphic–linguistic distinction described in the literature as a means to extend the graphic-linguistic distinction into aural domains. The model makes the distinction in terms of lower level perceptual capabilities that enable perceivers to perceptually process concrete structures (e.g., geometric and topological features) on the one hand, and higher level capabilities that enable perceivers to process and interpret how those perceptually processed structures fall under more abstract conceptual categories on the other.

Due to these distinctions, the model predicts that perceptual relations (conveyed via graphics or non-linguistic sonification) afford the communication of concrete relations (conveyed via text or text-to-speech) more effectively compared to conceptual relations conveyed via text or text-to-speech. In addition, the model predicts that conceptual relations (conveyed via text or text-to-speech) afford the

communication of abstract relations more effectively compared to perceptual relations conveyed via graphics or non-linguistic sonification.

In addition, the model streamlines accounts that distinguish diagrammatic from sentential structures to (1) characterize sentential structures as composed of conceptual relations among conceptual objects on the one hand, and (2) diagrammatic structures as perceptually represented relations among conceptual objects on the other. Under this account, (3) a sonic diagram is conceptualized as sonically conveyed relations among linguistically conveyed (via text-to-speech) objects.

This model may be applied by researchers and designers to generate testable predictions. For example comparing recognition performance of a visual data set communicated either through sonification or text-description. This is useful within a design context because designers lack guidelines for converting visual graphics into non-visual perceptual modes.

Acknowledgements

References

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral & Brain Sciences, 22,* 577–660.

Barsalou, L. W. (2003). Abstraction in perceptual symbol systems. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 358*(1435), 1177–1187. doi:10.1098/rstb.2003.1319

Barsalou, L. W. (2005). Abstraction as dynamic interpretation in perceptual symbol systems. In L. Gershkoff-Stowe & D. Rakison (Eds.), *Carnegie Symposium Series: Building object categories* (pp. 389–431). Majwah, NJ: Erlbaum.

Barsalou, L. W. (2009). Simulation, situated conceptualization, and prediction. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 364*(1521): 1281–1289. doi:10.1098/rstb.2008.0319

Barwise, J., & Etchemendy, J. (1990). Visual information and valid reasoning. In W. Zimmerman (Ed.), *Visualization in mathematics* (pp. 8–23). Washington, DC: Mathematical Association of America.

Brown, L. M., Brewster, S. A., Ramloll, S. A., Burton, R., & Riedel, B. (2003). Design guidelines for audio presentation of graphs and tables. *International Conference on Auditory Display.*

Coppin, P. W. (2014). *Perceptual-cognitive properties of pictures, diagrams, and sentences: Toward a science of visual information design* (Doctoral dissertation, University of Toronto, Toronto, Canada). Retrieved from https://tspace.library.utoronto.ca/handle/1807/44108

Damasio, A. R. (1989). The brain binds entities and events by multiregional activation from convergence zones. *Neural Computation, 1*(1), 123–132.

Gibson, J. J. (1986). *The ecological approach to visual perception.* Hillsdale, NJ: Lawrence Erlbaum.

Goodman, N. (1968). *Languages of art: An approach to a theory of symbols.* Indianapolis, IN: Bobbs-Merrill Company.

Hermann, T., & Hunt, A. (2011). *The sonification handbook.* Berlin: Logos Verlag.

Kosslyn, S. M., Ganis, G., and Thompson, W. L. (2001). Neural foundations of imagery. *Nature Reviews Neuroscience, 2*(9), 635–642. doi:10.1038/35090055

Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science, 11,* 65–99. doi:10.1111/j.1551-6708.1987.tb00863.x

Loomis, J. M., Klatzky, R. L., & Giudice, N. A. (2013). Representing 3Dspace in working memory: Spatial images from vision, touch, hearing, and language. In S. Lacey & R. Lawson (Eds), *Multisensory Imagery: Theory & Applications* (pp. 131-156). New York: Springer.

Mandler, J. M. (2006). Categorization, development of. In *Encyclopedia of Cognitive Science.* doi:10.1002/0470018860.s00516

Mauney, B.S. & Walker, B.N. (2004, July). Creating functional and livable soundscapes for peripheral monitoring of dynamic data. In *10th International Conference on Auditory Display* (ICAD 2004). (pp. 6-9).

Nasir, T., & Roberts, J. C. (2007, June). Sonification of spatial data. In *13th International Conference on Auditory Display* (ICAD 2007) (pp. 112–119). ICAD.

Palmer, S. E. (1978). Fundamental aspects of cognitive representation. In E. Rosch & B. B. Llyod (Eds.) *Cognition and Categorization,* 259–303. Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Pylyshyn, Z. W. (1973). What the mind's eye tells the mind's brain: A critique of mental imagery. *Psychological Bulletin, 80*(1), 1.

Russell, B. (1923). Vagueness. *Australasian Journal of Psychology and Philosophy, 1*(2), 84–92. doi:10.1080/00048402308540623

Shimojima, A. (1999). The graphic-linguistic distinction: Exploring alternatives. *Artificial Intelligence Review, 13*(4), 313–335.

Simmons, W. K., & Barsalou, L. W. (2003). The similarity-in-topography principle: reconciling theories of conceptual deficits. *Cognitive neuropsychology, 20,* 451–486.

Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception, & Psychophysics, 73*(4), 971–995. doi:10.3758/s13414-010-0073-7

Zhao, H., Plaisant, C., Shneiderman, B., & Lazar, J. (2008). Data sonification for users with visual impairment: a case study with georeferenced data. *ACM Transactions on Computer-Human Interaction* (TOCHI), 15(1), 4.